UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

DEVELOPING NOVEL QUANTITATIVE IMAGING ANALYSIS SCHEMES BASED MACHINE LEARNING FOR CANCER RESEARCH

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for

Degree of

DOCTOR OF PHILOSOPHY

By

SEYEDEHNAFISEH MIRNIAHARIKANDEHEI

Norman, Oklahoma

2021

DEVELOPING NOVEL QUANTITATIVE IMAGING ANALYSIS SCHEMES BASED MACHINE LEARNING FOR CANCER RESEARCH

A DISSERTATION APPROVED FOR THE

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Bin Zheng, Chair

Dr. Hong Liu

Dr. Joseph P. Havlicek

Dr. Tomasz Przebinda

I dedicated my dissertation to my family,

Especially to my wonderful mom and dad,

And also

My amazing husband.

Acknowledgments

I would like to express my deepest gratitude to my research supervisor, Dr. Bin Zheng, for his overwhelming attitude to help all of his students. I am extremely grateful for his invaluable support, guidance, and insightful comments throughout my Ph.D. I could not imagine having a better mentor and advisor for my Ph.D. study. It was also a great privilege that he offered me to participate in several research projects and collaborating with different research groups working at different institutes such as OU Norman, OUHSC, and OSU.

Besides my advisor, I am tremendously fortunate to have committee members: Dr. Hong Liu, Dr. Joseph Havlicek, Dr. Tomasz Prezebinda. I appreciated their encouragement and support that motivate the progress of this project.

I want to thank my fellow colleagues and friends, Dr. Faranak Aghaei, Morteza Heidari, Gopichandh Danala, and honor our colleague's memory, Dr. Yunzhi Wang.

I am extremely thankful to my parents for their love, prayers, and sacrifices for educating me and preparing me for my future.

Last but not least, I am forever indebted to my lovely husband for his understanding, endless patience, standing up for me, and supporting me.

Table of Contents

Contents

List of Tables

List of Figures

Abstract

The computer-aided detection (CAD) scheme is a developing technology in the medical imaging field, and it attracted extensive research interest in recent years. In this dissertation, I investigated the feasibility of developing several new novel CAD schemes for different cancer research purposes. First, I investigated the feasibility of identifying a new quantitative imaging marker based on false-positives generated by a computer-aided detection (CAD) scheme to predict short-term breast cancer risk. For this study, an existing CAD scheme was applied "as is" to process each image. From CAD-generated results, some detection features were computed from each image. Two logistic regression models were then trained and tested using a leave-one-case-out cross-validation method to predict each testing case's likelihood of being positive in the next subsequent screening. This study demonstrated that CAD-generated false-positives contain valuable information to predict short-term breast cancer risk. Second, I identified and applied quantitative imaging features computed from ultrasound images of athymic nude mice to predict tumor response to treatment at an early stage. For this study, a CAD scheme was developed to perform tumor segmentation and image feature analysis. The study demonstrated the feasibility of extracting quantitative image features from the ultrasound images taken at an early treatment stage to predict tumor response to therapies. Last, I optimized a machine learning model for predicting peritoneal metastasis in gastric cancer. For this purpose, I have developed a CAD scheme to segment the tumor volume and extract quantitative image features automatically. Then, I reduced the dimensionality of features with a new method named random projection to optimize the model's performance. Finally, the gradient boosting machine model was applied along with a synthetic minority oversampling technique to predict peritoneal metastasis risk. Results suggested

that the random projection method yielded promising results in improving the accuracy performance in peritoneal metastasis prediction.

In summary, in my Ph.D. studies, I have investigated and tested several innovative approaches to develop different CAD schemes and identify quantitative imaging markers with high discriminatory power in various cancer research applications. Study results demonstrated the feasibility of applying CAD technology to several new application fields, which can help radiologists and gynecologists improve accuracy and consistency in disease diagnosis and prognosis assessment of using the medical image.

Chapter1: Introduction

1.1 Concept and Steps of Computer-aided Detection (CAD) of Medical Images

Primary studies on computers' quantitative analysis of medical images were carried out in the 1960s [1-3]. At that point, it was generally believed that computers could substitute radiologists in identifying different abnormalities due to the fact that computers are better than human beings at performing particular tasks [4]. Therefore, the concept of computer diagnosis was formed at that time. Although researchers achieved impressive results from the initial studies, the first attempts to replace radiologists with computers were unsuccessful. The critical reasons for this failure were that the advanced image processing techniques and powerful computers were not available at that time. A few years later, another approach began in the 1980s, in which the computers' outputs were employed to help radiologists not replace them [4]. Currently, this concept is recognized as computer-aided detection and diagnosis (CAD), which has spread extensively.

To help develop and establish a new paradigm of precision or personalized medicine, the development and assessment of CAD schemes have been attracting broad research interest in the medical imaging informatics field, which aims to provide clinicians (i.e., radiologists, oncologists, and pathologists) the second opinion or decision-making supporting tools to more accurately detect, diagnose and/or treat diseases (i.e., cancers) in the clinical practice [5].

Generally, a typical CAD scheme consists of segmented regions, abnormality detections, and the extraction of their unique features for subsequent classification of the problem. CAD schemes' studies have contributed to the automation of these steps, either through improving an existing technique or developing a new technique. Several studies have been conducted to improve the

CAD system's performance by optimizing its every stage performance. In the following subsections, each step of the CAD scheme was described briefly.

1.1.1 Lesion or ROI Segmentation

Enhancing the accuracy of the segmentation for the region of interest (ROI) in the current CAD is one of the hot topics that attract many research interests. Available CAD schemes suffer from a technically perfect segmentation step, particularly when the system is trained on a large dataset. On the one hand, the segmentation step in the CAD systems cannot be fully automatic since it might not be accurate enough to segment all cases in a large dataset correctly. On the other hand, manual segmentation would not be the right choice for this kind of data set as it is time-consuming. Additionally, if segmentation is not accurate enough, features extracted from the segmented area, locally or globally, might not represent the lesions or ROIs. Hence, it would be vital to enhancing the accuracy of this step. In paper [6], the authors tried to improve the CAD scheme's segmentation and evaluate the performance in predicting short-term breast cancer risk. In this study, the authors used bilateral mammographic density asymmetry features of right and left breasts in a large dataset. They also proposed a new method for the preprocessing segmentation step. After applying the new method, the AUC value increased from 0.63 to 0.70 (**Figure 1-1** illustrated the comparison of the AUC value of the CAD system before and after applying the proposed segmentation step). The results of this study demonstrated the importance of investigating a better segmentation method affect the accuracy of the model significantly.

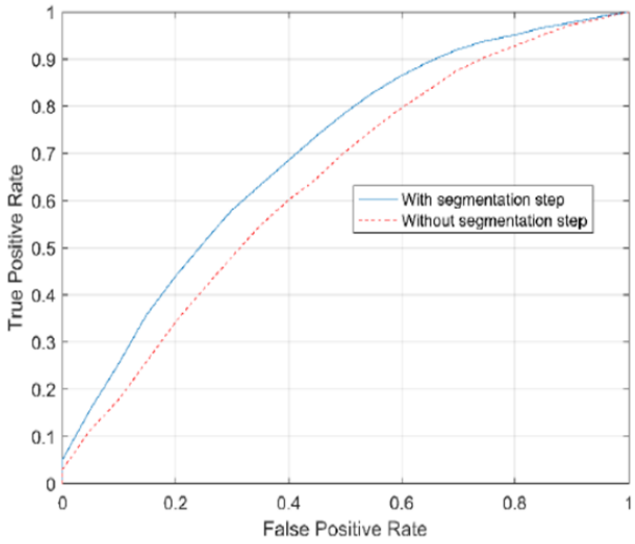


Figure 1-1. ROC curve of cancer risk model with and without the proposed segmentation step[6].

1.1.2 **Feature Extraction**

Once tumor regions are defined, imaging features can be extracted. Feature extraction means using the most efficient features which can represent valuable information for classification and analysis. Therefore, this step plays a critical role in improving the CAD schemes' performance, so that numerous studies have been carried out in this area to date.

Previous CAD mainly used morphological features such as lesion size and shape factors (i.e., circularity and boundary speculation), texture patterns (e.g., homogeneous, or heterogeneous), as well as features of tumor location and the surrounding tissues (e.g., near the heart) in 2D space [7]. While in recent studies, image features of a neighboring voxel in 3D space were introduced. These features describe autocorrelation, contrast, correlation, cluster prominence, cluster shade, cluster tendency, dissimilarity, energy, homogeneity, maximum probability, the sum of squares, sum average, sum variance, sum entropy, or difference entropy. Besides, gray level run length features, derived from run-length matrices. From the gray level run length matrix, features can be extracted

describing short and long-run emphasis, gray level nonuniformity, run-length nonuniformity, run percentage, low gray level run emphasis, and high gray level run emphasis [7].

Radiomics is an emerging field that transforms imaging data into a high-dimensional feature space through many automatically extracted data-characterization algorithms. In the paper [8], it was hypothesized that radiomic features capture distinct phenotypic differences of tumors and may have prognostic power and clinical significance across different diseases. In that study, 440 radiomic image features were defined that describe tumor characteristics and can be extracted in an automated way. The features were divided into four groups: (I) tumor intensity, (II) shape, (III) texture, and (IV) wavelet features. The first group quantified tumor intensity characteristics using first-order statistics, calculated from the histogram of all tumor voxel intensity values. The second group contained features based on the tumor's shape (sphericity or compactness of the tumor). The third group included textual features that are able to quantify intra-tumor heterogeneity differences in the texture that is observable within the tumor volume. These features were calculated in all three-dimensional directions within the tumor volume. Thus, each voxel's spatial location compared to the surrounding voxels was taken into account. Forth group features including the intensity and textural features extracted from wavelet decompositions of the initial image, thereby concentrating the features on different frequency ranges within the tumor volume. Then, feature selection based on stability ranks and performance was applied. Four radiomics signature, namely (I) ‘Statistics Energy’ describing the overall density of the tumor volume, (II) ‘Shape Compactness’ quantifying how compact the tumor shape is, (III) ‘Grey Level Nonuniformity’ a measure for intratumor heterogeneity and (IV) wavelet ‘Grey Level Nonuniformity HLH, also describing intratumor heterogeneity after decomposing the image in midfrequency, were selected

[8]. This study illustrated that applying radiomic features provides an excellent opportunity to advance decision-support in cancer treatment.

1.1.3 **Optimal Feature Selection**

Since researchers initially use CAD to compute a large number of image features, many of them can be redundant (highly correlated) or irrelevant (with lower performance), selecting a small set of optimal features is very important. For illustration, in paper [9], the authors investigated a new approach to improve feature selection efficiency in developing a CAD scheme. For this study, after segmentation of each suspicious lesions, 271 features were extracted categorized in different groups consisting of shape, texture, contrast, local topological features, isodensity, speculation, the features related to the presence and location of calcifications and fat, as well as texture features from the dilated region segments. To select optimal features from this initial large feature pool and build an efficient classifier, the authors explored and compared four different feature selection methods to optimize an artificial neural network (ANN) based classifier, namely: (1) Phased Searching, (2) Sequential Floating Forward Selection (SFFS) method, (3) Genetic Algorithm (GA), and (4) Sequential Forward Selection (SFS) method. Results of the four selected methods were assessed using a ten-fold cross-validation method. As it can be seen in **Table 1-1** among these four methods, SFFS has the highest efficiency, which takes 3%–5% of computational time and the highest AUC value.

This study revealed that using a new SFFS based approach significantly improves image feature selection efficiency for developing CAD schemes. More importantly, the results of this study suggested that choosing an appropriate feature selection method can notably affect the machine learning model's performance.

Table 1-1. Average AUC results and computation times per fold and corresponding standard deviation intervals for the four-compared feature selection and classification approaches computed over the tenfold cross-validation experiments.

No.	Method	AUC	Average Computation Time per Fold (min)
1	Phased Searching	0.856 ± 0.029	134.6 ± 3.6
2	SFFS	0.864±0.034	34.6 ± 1.2
3	GA_ANN	0.863 ± 0.036	1009 ± 12.4
4	SFS	0.835±0.02	13.3 ± 1.6

1.1.4 **Machine Learning**

In medical imaging, machine learning is a technique for identifying patterns and computing the image markers, which led to making predictions, detections, risk assessments, and diagnoses of interest [10]. The machine learning algorithm then identifies the best combination of the extracted image features for the image classifications or computations of some metric for the region of interest. Numerous machine learning methods can be used in the medical imaging field, each with different characteristics and applications [10-13]. Recently, convolutional neural networks (CNN) and deep learning achieved outstanding results in almost all computer vision areas. The great advantage of using deep learning is that it does not need identification and extraction of image features step; image features are identified in the learning process [10]. Machine learning has been used widely in the medical imaging field and will substantially influence the future. Computer-aided detection performed using a machine learning algorithm can

help clinicians interpret medical imaging findings and reduce interpretation time. As previously discussed, machine learning uses image properties and computing features for classifications. A recent study [22] authors declared that feature computation is one of the most critical steps of a machine learning technique and can highly affect a selected classification model's accuracy. To test the hypothesis mentioned above, the authors evaluate different classifiers' performance after applying other feature extraction methods. The results revealed that for specific feature extraction, the impact of choosing a classifier is not substantial. However, for a particular classifier selecting a proper feature extraction and computation method significantly affects the system's performance.

Hence, the main focus of all my studies during my Ph.D. was focused on investigating new image features, which are considered one of the essential steps of developing a machine learning algorithm.

1.2 Application and Advantages of CAD Schemes

Over the years, most published CAD research was focused on detecting and diagnosing different abnormalities in three organs such as chest, breast, and colon [4]. However, other organs' abnormalities, including liver, brain, and vascular systems, were also subjected to the CAD studies [4]. **Figure 1-2**. depicted 98 developed CAD schemes categorized based on their applications of detecting various abnormalities in the recent studies. According to the current research, CAD schemes have been developed mostly for detecting or diagnosis breast and lung cancer [14].

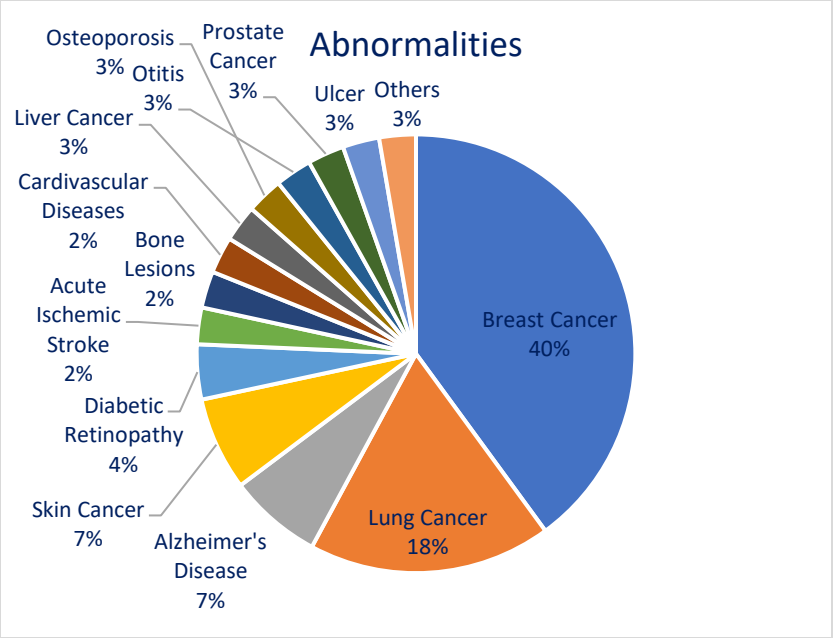


Figure 1-2. Categorized 98 CAD schemes in recent studies based on their applications in detecting different abnormalities.

Additionally, **Figure 1-3** shows the number of CAD schemes used for each imaging modality among the 98 CAD mentioned above schemes. The results demonstrated that in most studies, CT and mammography were used to develop the CAD schemes [14].

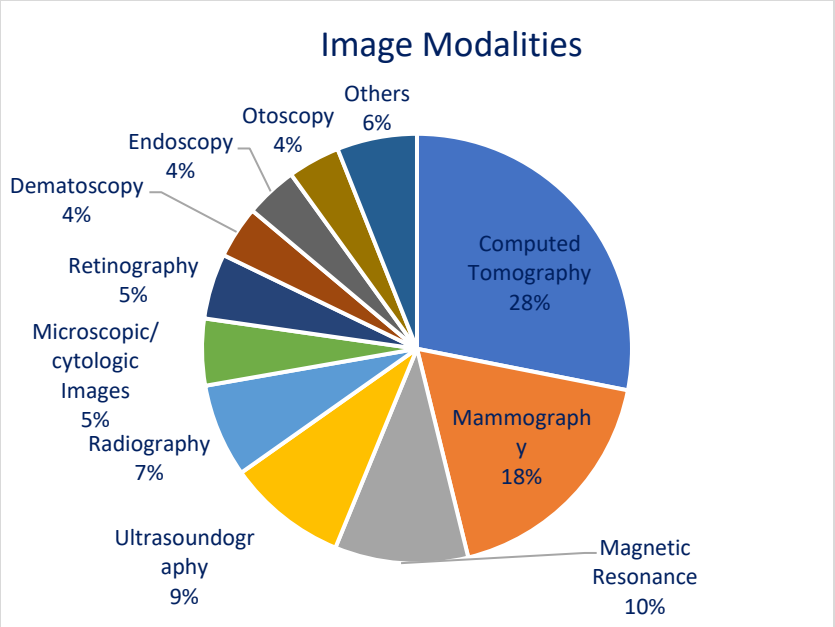


Figure 1-3. The number of CAD schemes used for different imaging modalities.

In mammography, several studies [15-19] conducting on large datasets ranging from 8,682 to 115,571 reported that the breast cancer detection rate improved after applying CAD schemes. Table 1-2 shows the summary of the results reported in mentioned studies. Moreover, in [19], authors noticed that after applying the CAD scheme for small invasive cancers, the detection rate increased about 64%. These results demonstrate the importance of applying CAD schemes to assist radiologists in detecting different abnormalities such as breast cancer.

Furthermore, in the study [20], the authors compared radiologists' accuracy performance with and without using the CAD scheme for detecting benign and malignant lung nodules on chest radiographs. As shown in **Figure 1-4,** the radiologists' performance was substantially improved after using CAD schemes [20].

Table 1-2. Summary CAD schemes effect in breast cancer detection in previous studies.

Authors	Dataset	Enhance in Recall Rate (%)	Enhance in Cancer detection rate (%)
Birdwell et al.[17]	8,682	7.6	7.4
Freer et al.[15]	12,860	18.8	19.5
Morton et al.[18]	18,096	10.8	7.6
Cupples et al.[19]	27,274	8.1	16.1
Gur et al.[16]	115,571	0.1	1.7

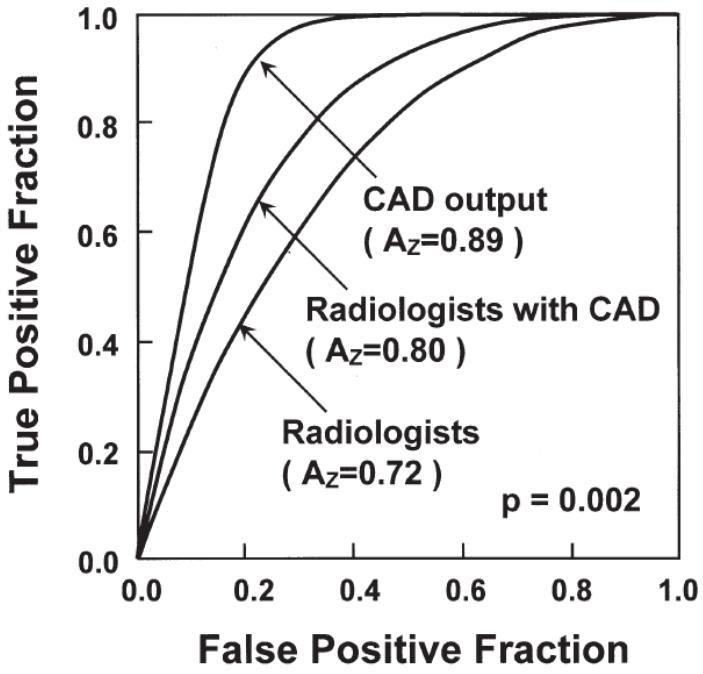


Figure 1-4. the ROC curves comparing radiologists' performance with and without the CAD scheme for detecting chest radiography abnormalities.

The applications of CAD schemes have also been extended to numerous image analysis quantitative techniques, including automatic lesion segmentation, lesion area measurement, and

dynamic flow information. These functions may also help clinicians and radiologists with the diagnostic examination, treatment set up, treatment response assessment, prognostic, and risk prediction for several diseases using image-based features individually or in combination with clinical features. After all, CAD schemes for assisting early disease diagnosis in different areas can improve treatment therapies and affect higher rates of positive outcomes.

1.3 Limitations of The Current CAD Development

Although, in recent years, researchers achieved outstanding results in developing CAD schemes, the efficiency of these schemes in helping physicians to improve the detection and diagnosis is still a controversial issue. Some studies alleged that the CAD scheme's high false-positive detection rates substantially reduce the radiologists' efficiency and accuracy [21-23] (**Figure 1-5**). On the other hand, some studies illustrated that radiologists could detect more occult cancer if they used the CAD scheme's assistance appropriately [24, 25]. Moreover, a recent study revealed that the observers' performance to detect subtle abnormalities was substantially affected by the CAD system's performance. The higher the performance of the cueing system, the more significant the improvement in the observer's performance.

Further, a low-performance cueing system could adversely impact the radiologists' performance to detect suspicious lesions. To clarify, the CAD systems with high false-positive results reduce the accuracy of the physicians' performance, while cueing systems with low false-positive results improve the radiologists' performance in detecting abnormalities. Also, the type of lesion, mass or microcalcification, directly affected the performance in detecting breast abnormalities.

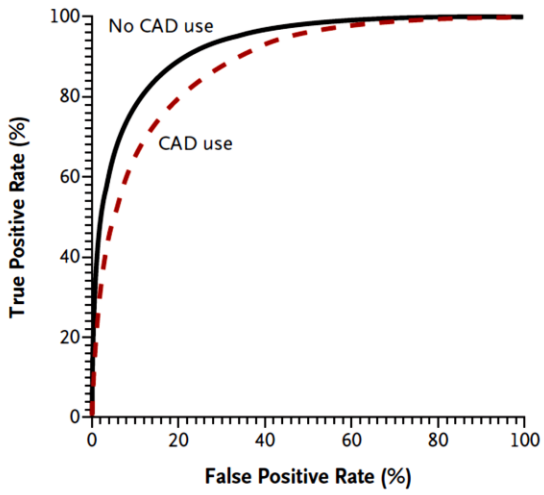


Figure 1-5. The overall accuracy of screening mammography, according to the use of CAD systems.[23]

In addition to the limitations mentioned above, in the following paragraphs, I have mentioned some of the significant challenges and difficulties in developing CAD schemes in current studies:

1. One of the main challenges encountered in CAD schemes' development is that the new methods' results can vary depending on the used datasets. Hence, it is not an easy task to determine the effectiveness of a new technique. To establish a technique's reliability, tests should be performed with a set of images that preferably have various acquisition characteristics.

2. Another big issue regarding the development of CAD systems is the insufficient availability of test images. It is not always viable to access a medical image database containing various acquisition characteristics, structures, or abnormalities that a technique requires the detection, analysis, or diagnosis of a disease.

3. Precise Segmentation of suspicious lesions or regions of interest (ROIs) is also a challenge for researchers in CAD schemes' development. The accuracy of lesion or ROI segmentation can affect the accuracy of feature computation and performance of the final CAD-generated detection or classification scores. Hence, it is an essential step in developing the CAD scheme; however, due to the significant heterogeneity of lesions and tissue background, the automatic segmentation of lesions or ROIs is still a difficult task.

4. Optimal and efficient imaging-based risk factors identified from medical images with CAD schemes have been another difficulty of the current CAD schemes. Currently, many different image-based risk factor has introduced in several studies. However, improving the CAD schemes' accuracy performance by investigating the most efficient risk factor is still controversial.

5. Extracting image markers with high discriminatory power to detect and predict different abnormalities is counted as another difficulty of current CAD schemes in recent studies since the extracted features from medical images can significantly affect the CAD scheme's performance.

6. In medical images, a large number of image features are calculated from the segmented areas (i.e., based on the Radiomics concept). Nevertheless, many of the calculated features do not have much discriminatory power and are redundant. Hence, identifying and selecting an optimal small set of image features is another challenge facing the performance improvement of CAD schemes.

7. Conventional CAD schemes use a "black-box" type approach, which reduces the users' confidence (clinicians) to consider and accept CAD-generated results. To overcome this

issue, it is vital to increase transparency to better understand CAD schemes' reasoning and develop interactive CAD tools (or graphic user interface).

8. Since medical imaging technology is fast-changing, the current paradigm for developing CAD systems cannot keep pace with these changes, specifically, as new imaging systems are established, where clinical images are insufficient. The ultimate purpose should be developing the field to the stage where it will be feasible to model the imaging system's characteristics and model clinicians' performance. Then, from an array of image processing, pattern recognition techniques, and artificial intelligence, choosing a group of methods leads to the optimal CAD system.

Despite the limitations and difficulties of developing and applying CAD schemes in clinical practice, many researchers believe that CAD has great potential in future clinical applications, and improve its performance and application methods is needed [26]. Thus, CAD remains a very active research topic in the medical imaging field, and many advanced or improved technologies have been investigated and applied in CAD research.

Chapter 2: Research Objectives and Hypotheses

2.1 Overall Objectives of Research in This Dissertation

Previous studies have provided ample scientific data and evidence to support that CAD of medical images has broad application potential to provide clinicians (i.e., radiologists and oncologists) quantitative decision-making supporting tools for more accurately detecting cancer or more effective treating cancer patients. The comprehensive literature search and review during my Ph.D. studies have helped me better understand the advantages, progress, and challenges of developing and applying CAD schemes of medical images to add more value in various clinical application fields. Specifically, I recognize that to make a new and significant contribution in the CAD research field, researchers should continue to explore new research ideas for identifying application approaches and developing or using advanced image processing and machine learning technology to improve and optimize CAD schemes. Thus, the overall objective of my research in this Ph.D. dissertation is to explore and investigate three new CAD applications using the novel or unique research ideas and/or the advanced image feature processing technology, which include (1) identify a new novel CAD-generated risk model to predict short-term breast cancer risk for improving the efficacy of breast cancer screening, (2) develop and analyze a quantitative imaging marker of ultrasound images to help evaluate and select optimal thermal therapy drugs and methods, and (3) investigate and apply a novel random projection algorithm to optimize machine learning models for better stratify gastric cancer patients for taking neoadjuvant chemotherapy to improve treatment efficacy. The details regarding the motivations and hypotheses of conducting studies of the above three new CAD applications are discussed in the following sections of this Chapter.

2.2 Applying A CAD-Generated Image Marker to Predict Breast Cancer Risk

2.2.1 Background and Motivation

Breast cancer causes the highest mortality rate among young women in the United States [27]. Previous studies revealed that early detection plays a critical role in reducing the mortality rates due to breast cancer [27]. For early detection, mammography is one of the most efficient imaging modalities for breast cancer screening [27]. However, the efficacy of current mammogram screening is controversial since cancer-detecting yield in the mammography screening environment is low [28]. In addition, it also generates high false-positive results. Hence, risk-based screening or more accurate prediction models are required to help radiologists better interpret the mammography results. Therefore, investigating a better or more accurate and useful risk factor makes a considerable contribution to this field, which was my first study's motivation.

2.2.2 Hypothesis and Scientific Rationale

Although CAD schemes have been widely used in clinical practices to process screen mammograms, they still generate high false-positive results, which reduce radiologists' specificity level [29]. In the first study, I hypothesized that we could use these high false-positive results as a risk factor; based on the previous studies' findings, which I discuss in the following paragraph.

Although there are many different risk factors in medical imaging, mammographic density is considered a high cancer risk factor in current models [30]. Authors in previous studies declared that high mammographic density would lead to high false-positive recalls [31]. On the other hand, although a higher false-positive recall rate is one of the significant hurdles that significantly reduce the efficacy of mammography screening, previous studies have shown that the false-positive recall might be a potential risk factor to indicate the risk or higher probability of cancer detection in the

subsequent mammography screening [32, 33]. Thus, in the first study, I hypothesized that CAD-generated false-positives might not be harmful or only had a negative impact, which may contain valuable information.

2.3 Applying CAD Method to Predict Efficacy of Cancer Treatment

2.3.1 Background and Motivation

Currently, imaging plays an essential or determinative role to assess tumor response to the treatment, such as using the RECIST guideline to measure the change in tumor size before and after treatment. Besides, ultrasound is a low-cost, portable and easy-to-use image modality, which has recently emerged as a promising modality for imaging, quantitatively monitoring tissue changes, and drug effect over a specific time frame [34]. Using ultrasound imaging modalities to identify treatment effects on cancerous cells has been attracting extensive research interest in recent years. For this purpose, identification and optimal selection of new biomarkers that have high discriminatory power to predict treatments' results play a crucial role. Thus, the second study's motivation was to compute quantitative image features from ultrasound images, which may provide useful imaging markers to predict tumor response to treatment or identify more effective treatment methods.

2.3.2 Hypothesis and Scientific Rationale

In the second study, I hypothesized that by extracting new quantitative image features from ultrasound images, we are able to predict the efficacy of cancer treatment at an early stage. In the following, I mention the scientific grounds for this hypothesis.

In previous studies, the efficacy of different treatments has been investigated by extracting quantitative features from other imaging modalities such as CT, MRI, X-ray images [35-38]. Those

studies have shown that extracting imaging features from those modalities assist clinicians in predicting the cancer treatment's efficacy. On the other hand, recently, ultrasound has been widely used to evaluate the treatments' response to cancerous tumors. However, whether or not the extracted features from the ultrasound images can provide any valuable information to assist in predicting the efficacy of treatment has not been investigated before my study. Therefore, in the second study, I have examined the feasibility of extracting quantitative imaging markers from ultrasound images to predict the treatments' efficacy on cancerous tumors.

2.4 Exploring New Method to Generate Optimal Features and Improve CAD Performance

2.4.1 Background and Motivation

Some studies revealed that novel radiomics techniques could extract quantitative information from medical images and assist in image interpretation [7, 39-48]. However, CAD schemes are initially computing a large number of features, and most of them are highly correlated with lower performance. Hence, selecting a small set of features to reduce the feature dimension and enhance learning accuracy is fundamental. My third study's motivation was to investigate a new feature dimensionality reduction method with higher accuracy performance than the conventional methods in the medical imaging field.

2.4.2 Hypothesis and Scientific Rationale

Optimal feature dimensionality reduction is essential in improving the performance of the machine learning model. In the third study, I hypothesized that applying a random projection (RP) feature reduction method would substantially enhance the accuracy performance of machine

learning model. The scientific foundation of this hypothesis has been discussed briefly in the following paragraph.

Previously the RP method has been applied in many different image processing tasks and computer vision applications [49-51]. Although it has not been used in the medical imaging field before my study, it showed promising results in other areas in reducing the features' dimensionality and enhance the accuracy of different models. Therefore, in the third study, I assumed that this method could be applied to the medical imaging field and improve the machine learning model's performance.

2.5 Organization of The Dissertation

This dissertation reports three of my main studies of developing quantitative medical imaging analysis that apply different machine learning algorithms (Chapter. 3,4,5,6). In Chapter 3, I have conducted a study to take advantage of the false positive generated by the CAD scheme to predict short-term breast cancer risk. In Chapter 4, I developed a quantitative ultrasound image feature analysis scheme to assess tumor treatment efficacy using a mouse model. In Chapter 5, a random projection algorithm was applied to optimize a machine learning model for peritoneal metastasis prediction in gastric cancer patients using CT images. Last, in Chapter 6, a summary of these new CAD schemes is discussed, which generates this dissertation's conclusion, including this candidate's future work.

Chapter 3: Exploring A New Imaging Marker Generated By Computer-Aided Detection Scheme To Predict Short-Term Breast Cancer Risk

3.1 Introduction

Mammography is one of the most commonly used imaging modalities in population-based breast cancer screening to date. However, the efficiency of screening mammography is controversial [52] because of the comparatively lower detection sensitivity (particularly among women with dense breasts or younger than 50 years old) [53] and higher false-positive recall rates [54] with the probable long-term psychosocial consequences [55]. To enhance breast cancer screening effectiveness, developing a new and more efficient personalized breast cancer screening paradigm has been recently attracting extensive research interest [56]. A prerequisite for realizing an optimal personalized screening is recognizing more effective breast cancer risk factors or developing more precise risk prediction models to stratify women into two groups with a higher and lower risk of having or developing breast cancer in the short-term (i.e., < 1 to 3 years). Accordingly, using these risk factors or prediction models may help clinicians or the individual patient better decide whether she should currently have more recurrent screening (e.g., annually) or be screened at longer intervals until her short-term risk considerably increases in future reassessments.

Although many epidemiological studies based on breast cancer risk prediction models have been previously developed and applied to detect high-risk women [30], the models mainly predict long-term (or lifetime) risk of subgroups of women comparing to the overall population, which have little discriminatory power to determine who should or should not be screened in the short-

term to help improve the efficiency of breast cancer screening [57]. Consequently, exploring new breast cancer risk factors and investigating new breast cancer risk prediction models remains an arduous task [58], attracting great research interest and effort. In this field, many researchers believe that mammograms consist of significant phenotype markers that can be quantified and used to improve the prediction of breast cancer risk [59]. In our previous studies, we investigated a new quantitative imaging marker based on the extracted bilateral asymmetry of mammographic density features between the left and right breasts. We showed an increasing trend in predicting short-term risk or the probability of women having or developing mammography-detectable cancers next following mammographic screening [60, 61].

During our studies of identifying new quantitative imaging features, I recently recognized a potential new method. In order to help radiologists reading and interpreting mammograms, computer-aided detection (CAD) schemes have been used as "the second reader" in a large number of breast imaging clinics since early 2000 [62]. Whether or not using CAD can assist enhance the accuracy of radiologists in breast cancer detection is also controversial, primarily because of the high false-positive detection rates generated by CAD schemes of mammograms [29]. Nevertheless, whether I can apply the quantitative image marker analysis and detection scores on the CAD-generated false-positive regions to assist short-term breast cancer risk prediction has not been investigated yet. Therefore, I proposed a new hypothesis. Since previous studies have indicated that CAD enabled identifying more early abnormalities that were either missed or overlooked by radiologists and later became image-detectable cancer [63, 64], the CAD-generated false-positives on the negative images may not be completely useless. As a matter of fact, the CAD-generated false-positives might contain useful information as quantitative imaging markers to assist short-term breast cancer risk prediction. The purpose of this study is to examine our

hypothesis using a relatively large and diverse image dataset of 1,044 negative mammography screening cases. After using a CAD scheme to process these negative images, I applied a machine learning model to explore a new quantitative imaging marker for risk prediction of having mammography-detectable cancer at the next subsequent screening (12 to 18 months later).

3.2 Materials and Method

3.2.1 Image Dataset

In this study, I utilized full-field digital mammography (FFDM) images chose from the pre-existing image dataset in our laboratory, which has been retrospectively assembled in our previous studies to develop CAD schemes of mammograms [65, 66] and quantitatively imaging markers for breast cancer risk prediction [60, 61]. Briefly, by excluding the interval cancer cases and the cases without four images of both the mediolateral oblique (MLO) and craniocaudal (CC) view of the right and left breasts, the dataset employed in this study includes FFDM images attained from 1,044 women who participated in routine annual mammography screenings. Each case had at least two or more subsequent FFDM screenings of four view images of both breasts. The latest screening is named “current” screening, which can be either positive or negative. All “prior” screenings are negative as reported by radiologists in the original image reading and interpretation, which may include a small fraction of “false-negative” cases in which the “early suspicious tumors” may be considered detectable in the retrospective review [65]. Nonetheless, since these “suspicious tumors” were either missed or overlooked by the radiologists in the initial screening, these cases' previous images are categorized as negative images in the clinical database. Therefore, the negative images used in this study are determined by the clinical record generated from the real mammography screening.

From this dataset, I chose images acquired from the first "prior" mammography screening, which was taken 12 to 18 months prior to the "current" mammography screening. In the "current" screening, 402 cases were positive in which cancer was detected from mammograms and confirmed by biopsy, while the rest of 642 cases remained negative (cancer-free) in the "current" screenings. Hence, these 1,044 "prior" negative screenings were categorized into two groups of the high and low-risk cases. **Table 3-1** summarizes the general information of the cases involved in our dataset.

Table 3-1. General information of the image dataset.

Risk Factor	Category	High- Risk Cases	Low-Risk Cases
Total Cases		402	642
Age (years old)	< 45	29 (7%)	169 (26%)
	45-55	93 (23%)	251 (39%)
	55– 65	148 (37%)	146 (23%)
	> 65	132 (33%)	76 (12%)
	Mean ± SD	60.88±11.10	52.48±10.52
	Median	60	50
Density BIRADS	Almost all fatty tissue	22 (5%)	40 (6%)
	Scattered fibroglandular densities	153 (38%)	248 (39%)
		217 (54%)	328 (51%)

	Heterogeneously dense	10 (2%)	26 (4%)
	Extremely dense		
	No family history was known		
Family History	Cancers in the 1st-degree relatives	149 (37%)	285 (44%)
	Cancers in the 2nd-degree relatives	48 (12%)	50 (9%)
	Cancers in the 3rd-degree relatives	59 (15%)	83 (13%)
	other	3 (1%)	22 (3%)
		143 (36%)	202 (31%)

3.2.2 Machine Learning Model

Next, as illustrated in the flowchart of **Figure 3-1**, the following steps are taken to process each image, extract CAD-generated features, and build a machine learning classifier to predict cancer risk. First, an existing CAD scheme of mammograms [67] was utilized. The previous study revealed that this CAD scheme's performance was quite comparable to two leading commercialized CAD schemes in identifying malignant breast lesions using an independent testing image dataset to all three CAD schemes under comparison [68]. In brief, the CAD scheme employs three stages to detect suspicious mass regions depicting on an image. First, a Gaussian bandpass

filter is applied to detect initially suspicious seeds (i.e., typically 10 to 50 per image depending on the complexity of breast tissue structure). Second, an adaptive multi-layer topographic region growing algorithm is applied to segment each suspicious region. Based on the set of region growing criteria, this step typically results in discarding more than 50% of suspicious regions initially detected in step one. Third, a multi-feature based artificial neural network (ANN) is applied to process each remaining suspicious region and generate a detection score, which indicates the likelihood of the detected region associating with a malignant lesion. In this study, I applied this CAD scheme of mammograms "***as is***" to process each FFDM image of all study cases in our dataset.

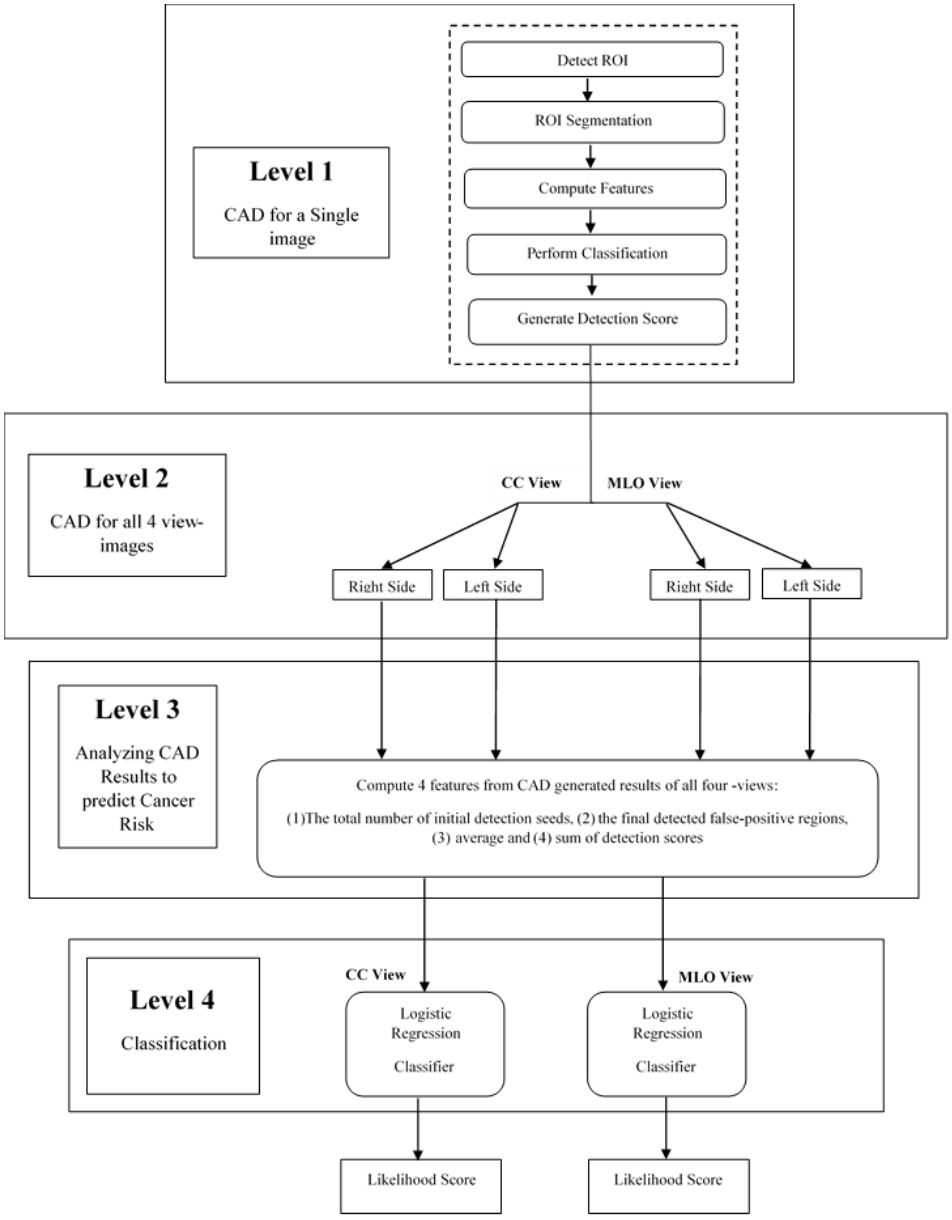


Figure 3-1. Flowchart showing the steps of applying the proposed image processing and risk prediction scheme.

Second, I summarized CAD-generated detection results and scores on all detected suspicious regions on each image. First, four features were computed, which are (1) the number of initial suspicious seeds detected by the first step of the CAD scheme, (2) the number of final suspicious mass regions detected by the second step of the CAD scheme, (3) total (summation) score of all

detected suspicious mass regions, and (4) average score of all detected suspicious mass regions. Second, I computed bilateral summation between two left and right (CC or MLO view) images. In this way, I generated a total of eight features, namely four for the CC view and four for the MLO view images. Third, I computed the mean value (μ) and standard deviation (σ) of each feature computed from all 1044 testing cases. Each feature was then normalized to a range from 0 to 1 (within $\mu \pm 2\sigma$). All outliers were assigned to either 0 or 1 depending on whether their actual value is smaller than zero or greater than one.

Third, since a mammogram is a two-dimensional projection image overlapped with fibro-glandular tissues along the projection direction, mammographic tissue patterns or image features computed from the CC and MLO view of one breast may vary significantly. Thus, I separately built two classifiers that combine the four bilateral summation features computed from either CC or MLO view images to stratify the testing cases into the high and low-risk classes of having mammography-detectable cancers in subsequent mammography screening. Despite the fact that many different machine learning classifiers can be used for this purpose, I selected a multinomial logistic regression model based classifier because it is inherently simple, low variance, fast in training, and has a lower probability of overfitting.

Fourth, in order to build the classifier, I used a Weka data mining and machine learning software platform [68], which has been successfully applied and tested in a number of our previous studies in developing multi-feature fusion based machine learning classifiers to predict cancer risk and prognosis [69-71]. In order to minimize the case selection or partition bias, I applied a leave-one-case-out (LOCO) based cross-validation method to train and test the classifier [72] in which a classifier was trained using 1,043 cases and tested using one remaining case. Thus, through 1,044 training and testing iterations for each classifier, each case (either two bilateral CC view or MLO

view images) had a CAD-generated independent classification score ranging from 0 to 1. The higher score indicates the higher risk or likelihood of the study case having or developing mammography-detectable cancer in a short-term, defined as occurring 12 to 16 months later.

Last, I conducted the following data analysis tasks to assess the performance of using the new risk prediction models for case stratification. I used a receiver operating characteristic (ROC) based data analysis method. A maximum likelihood-based ROC curve fitting program (ROCKIT, http://xray.bsd.uchicago.edu/krl/roc_soft.htm) was used to generate the ROC curve and compute the area under the curve (AUC). I also applied an operating threshold (T = 0.5) on the classification scores to divide the cases into two groups and generated the corresponding confusion matrix. I then computed the odds ratio (OR) from the confusion matrix. The adjusted odds ratios and their increasing trend were also computed and analyzed using a statistical software package (R version 2.1.1, http://www.r-project.org). The data analysis results were tabulated and compared.

3.3 Results

Figure 3-2 shows an example of applying our CAD scheme to process four bilateral CC and MLO view images of a "prior" mammography screening case. **Table 3-2** summarizes the computed four features and detection scores for all suspicious regions detected by the CAD scheme on four view images of this testing case. In this case, CAD initially detected 81 (ranging from 14 to 23) seeds for the suspicious lesions in the first step. After applying the second step of region growing algorithms, the suspicious mass numbers were reduced to 32 suspicious regions detected in four images. In the third step of the CAD scheme, an artificial neural network-based classifier generated a detection score for each of 32 regions, which indicates the likelihood of the region associated with a positive mass. Thus, unlike a conventional CAD cueing method used in the clinical practice, which only cues the regions with detection scores greater than a predetermined

threshold, all 32 suspicious regions are cued in **Figure 3-2,** and all detection scores reported in **Table 3-2** were extracted and used to build the prediction model. All 32 CAD-generated cueing markers were discarded as false-positives, and the case was classified as a negative case. The radiologists detected a malignant mass-type lesion in the subsequent mammography screening, as shown in **Figure 3-3**. Comparing images in **Figures 3-2** and **3-3**, I observed that the lesion detected in "current" images is not "visible" in the "prior" images. However, CAD has two cueing markers with detection scores of 0.56 and 0.65 (as shown in **Table 3-2**) on the left MLO view "prior" images (**Figure 3-2**), which seems to match with the region that has malignant mass detected in the "current" left MLO image (**Figure 3-3**).

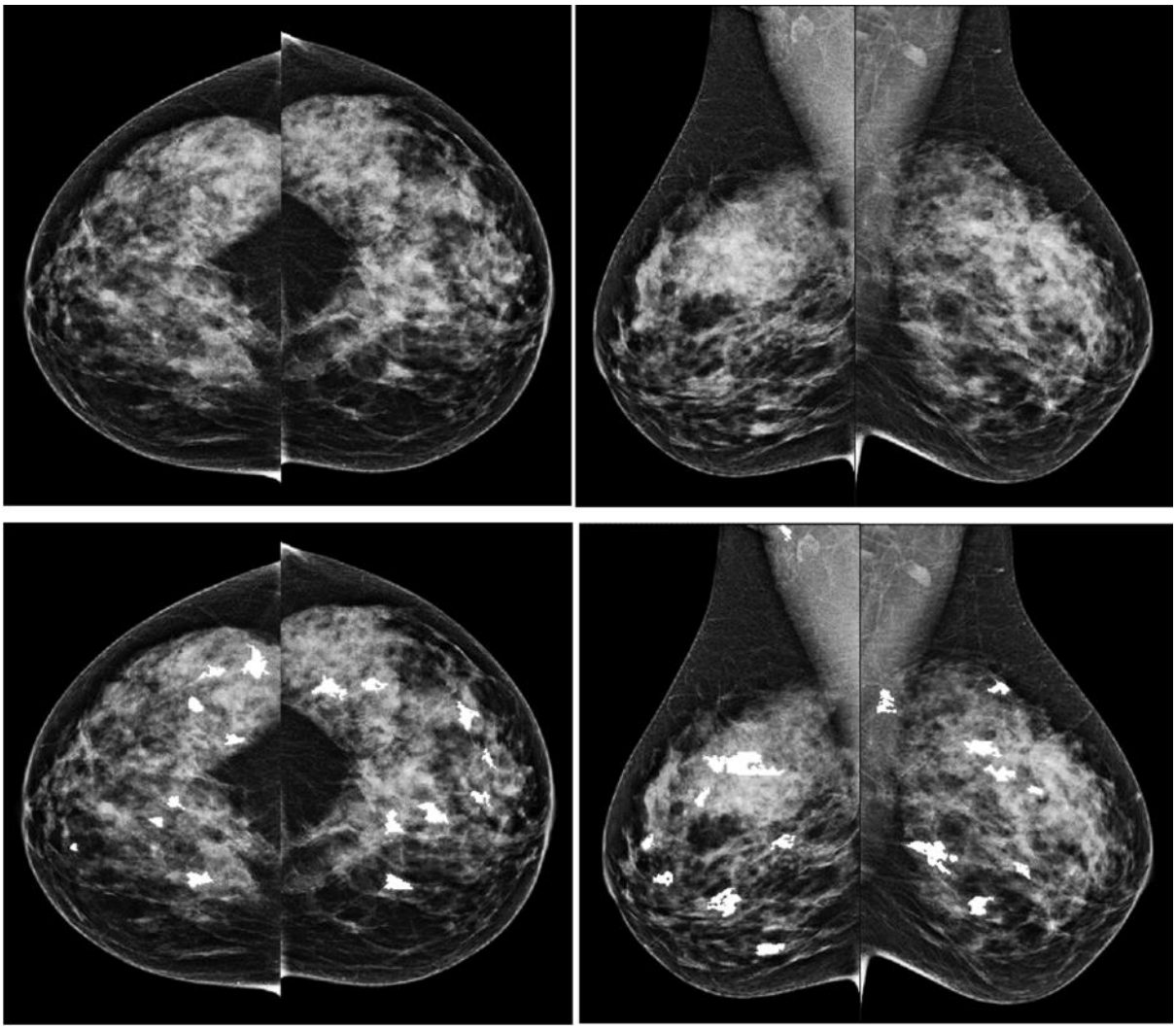


Figure 3-2. An example showing CAD processing results in one testing case of four bilateral CC and MLO view images. The top row shows four original images, and the bottom row shows images marked with CAD-detected suspicious mass regions. The CAD-generated detection scores.

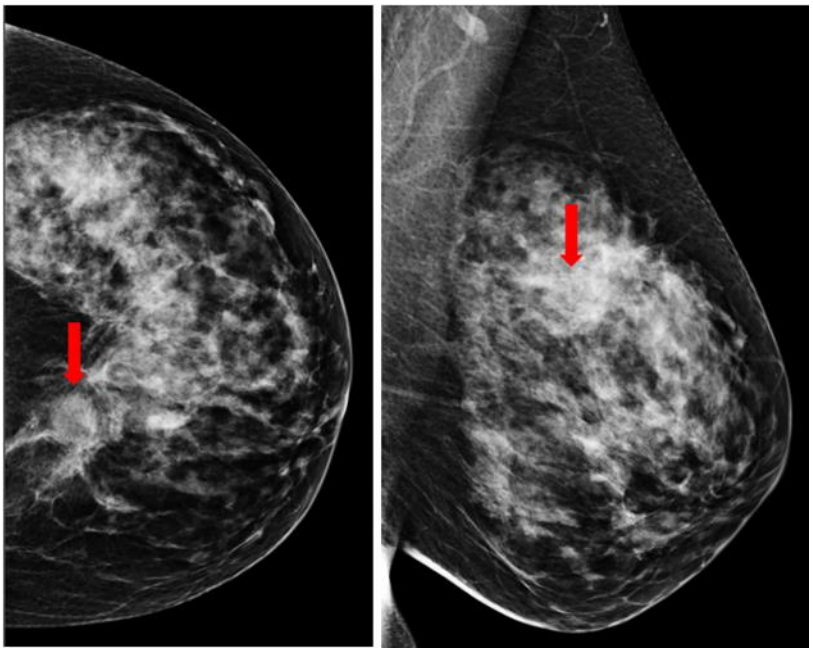


Figure 3-3. A malignant lesion was detected in the CC and MLO views of the left mammogram (as pointed by the arrow) in the next subsequent ("current") screening of the same case as shown in Figure 3-2.

Table 3-2. Summary of CAD-detection result and computed four features from one example of mammography screening case.

Image View	CC (left)	CC (right)	MLO (left)	MLO (right)
Detected regions	8	8	8	8
Average score	0.47	0.38	0.48	0.38
Total score	3.74	3.01	3.82	3.05
Initial suspicious seeds	23	14	22	22
Detailed detected regions and their scores	1) 0.44	1) 0.41	1) 0.27	1) 0.34
	2) 0.34	2) 0.39	2) 0.31	2) 0.31
	3) 0.50	3) 0.66	3) 0.56	3) 0.54
	4) 0.46	4) 0.27	4) 0.65	4) 0.19
	5) 0.32	5) 0.34	5) 0.38	5) 0.41
	6) 0.51	6) 0.31	6) 0.55	6) 0.26
	7) 0.52	7) 0.28	7) 0.40	7) 0.44
	8) 0.66	8) 0.37	8) 0.70	8) 0.58

Two logistic regression model based classifiers trained using CC and MLO view images yielded areas under ROC curves, AUC= 0.5861±0.0181 and AUC= 0.6521 ± 0.0176, to predict cancer risk, respectively. It shows that using CAD-generated detection results or features computed from bilateral MLO view images yielded higher prediction performance than using

bilateral CC view images ($p < 0.05$). **Figure 3-4** shows the ROC curve of applying the logistic regression model based classifier to MLO view images.

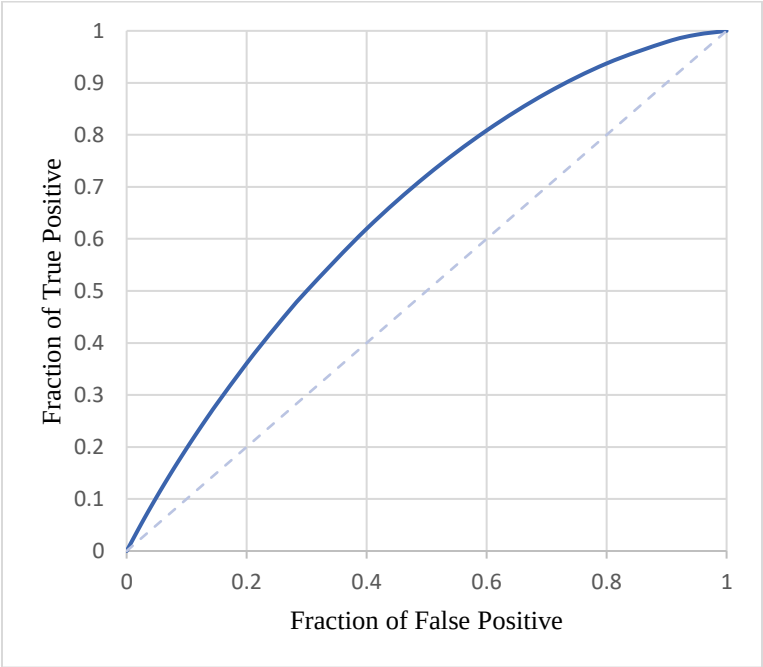


Figure 3-4. The ROC curve for MLO view with an AUC value of 0.6521 ±0.017.

Table 3-3 shows a confusion matrix of using CAD-generated detection scores on the bilateral MLO view images, generated by applying an operation threshold of T=0.5 on the risk prediction score. Using this operation score, the risk prediction sensitivity is 27.6%, and the specificity is 88.0%. **Table 3-4** summarizes the correspondingly computed odds ratio and risk ratio with their 95% confidence intervals. **Table 3-5** reports the adjusted odds ratios, which were computed after applying a set of thresholds to automatically divide 1,044 cases into five subgroups with an approximately equal case number. The risk prediction scores gradually increase from subgroup one to five. Using subgroup one as a reference (baseline) with 208 cases, the adjusted odds ratios and 95% confidence intervals for subgroups two to five with 209 cases were computed. The maximum adjusted odds in subgroup five's ratio is 4.49 with a 95% confidence interval of [2.95,

6.83]. Using regression analysis, the regression line's slope generated from the adjusted odds ratios is significantly different from the zero slopes ($p < 0.01$), which indicates an increasing trend of the adjusted odds ratios with an increase of risk prediction scores.

Table 3-3. The confusion matrix of using CAD generated detection results on bilateral MLO view images.

	Actual Positive	Actual Negative
Predicted Positive	**111**	77
Predicted Negative	291	**565**

Table 3-4. The odds Ratio and Risk Ratio of using CAD generated detection results on bilateral MLO view images.

Significance Level	95%
Odds Ratio	2.03 < **2.80** < 3.87
Critical Odds Ratio (COR)	1.11
Risk Ratio	1.77 < 2.30 < 2.99

Table 3-5. The adjusted Odds Ratio and 95% confidence interval using risk prediction scores computed from bilateral MLO view images.

Group Number	High-risk cases	Low-risk cases	Adjusted Odds Ratio	95% Confidence interval
1	51	157	1	N/A
2	55	154	1.09	0.71 – 1.71
3	80	129	1.91	1.25 – 2.91
4	92	117	2.42	1.59 – 3.68
5	124	85	**4.49**	2.95 – 6.83

3.4 Discussion

The impact of false positives on the subsequent risk of breast cancer has been previously investigated. For example, one study reported that women undergoing false-positive mammography at the first screening were less likely to participate in subsequent screenings, yet were more likely to develop interval cancers or cancers at subsequent screening [73]. Another study reported that women having false-positives involving fine-needle aspiration cytology or a biopsy had a higher breast cancer detection risk than those involving additional imaging procedures alone in subsequent screening participants over 17 years. The odds ratios ranged from 1.81 to 2.69 [33]. This study provided new evidence and experimental data regarding the possible association between false-positives and cancer detection risk in the subsequent screenings. Unlike

previous studies using qualitative assessment, our approach aims to explore, identify and/or develop a new quantitative imaging marker based on CAD-generated false-positive detection results to help predict short-term breast cancer risk.

Since CAD schemes of mammograms are currently available and used in clinical practice, extracting a new quantitative imaging marker from CAD findings in the "current" negative images is a more efficient and cost-effective approach to help predict the short-term risk of breast cancer detection in the subsequent screenings, which does not require additional imaging or genomics tests. Since the dataset used in this study was initially assembled for developing CAD schemes, it is not an age-matched dataset (as shown in **Table 3-1**) and thus includes a large fraction of difficult negative cases of younger women with dense breasts [60, 65]. Thus, applying the CAD scheme to this dataset tends to produce more false-positive detections, as indicated in the previous study [68]. However, despite using this relatively challenging dataset, the result of the risk prediction in this study is encouraging, which yielded an odds ratio of 2.80 (as shown in **Table 3-4**) or the maximum adjusted odds ratio of 4.49 (as shown in **Table 3-5**). The odds ratio is higher than or quite comparable to many existing risk factors reported in previous breast cancer risk prediction studies [30, 33, 73]. Thus, this study identified a potential new clinical application for CAD schemes of mammograms. The study also indicated that further exploration is worth optimizing this new CAD-based imaging marker in future studies.

In this study, I also had several specific observations. First, although CC and MLO view mammograms provide complementary information and reading mammograms of both views can help detect more cancers and reduce false-positive recalls [74], CAD results on two bilateral CC and MLO view images were highly correlated. In this dataset, CAD results on MLO images yielded higher risk prediction performance. AUC value for using CAD-generated results on MLO

view images was 0.652±0.017 with 95% CI of [0.617, 0.686] while using CAD results on CC view images, AUC value was 0.586±0.018 with 95% CI of [0.550, 0.621]. The computed correlation coefficients are 0.75 for all classification scores of 1,044 cases (or 0.76 and 0.73 for 402 high-risk and 642 low-risk cases, respectively). Thus, due to the higher correlation coefficient of CAD-generated detection results between using bilateral CC and MLO view images, using the conventional fusion methods [75, 76] was unable to further increase AUC value beyond that yielded using MLO view images. To clarify, using an average fusion method yielded AUC=0.631± 0.018.

Second, previous studies have reported that in the retrospective review of the positive cases' prior images, a large fraction of subtle or occult "early tumors" can be detected by radiologists. Thus, in developing the conventional CAD schemes of mammograms, these missed or overlooked "early tumors" by the radiologists in the real screening environment are typically considered "false negative" cases and selected as positive training cases to increase CAD sensitivity to detect more subtle tumors. In this study, I also conducted a test by removing 53 cases in which the "early masses" were considered "visible" or detectable in our previous retrospective review [65]. After removing these cases, the computed AUC value remained relatively constant with only a slight change (or increase) from 0.652±0.017 to 0.659±0.018. The results suggest that unlike the task to develop the conventional CAD schemes of mammograms, for this new specific task of predicting the risk of having mammography-detectable cancer in the subsequent (or annual) screening, the negative case group should include all cases that were determined as negative in the real mammography screening environment (include those potential "false negative" cases). More accurately predicting the risk of cancer detection in the next annual screening of these "false negative" cases can help prevent further cancer detection delay.

3.5 Conclusion

I investigated and demonstrated the feasibility of extracting a new quantitative mammographic imaging marker from the existing CAD-generated false-positives on the negative mammograms to help predict cancer detection risk in the next subsequent mammography screening. This new imaging marker or risk prediction model's discriminatory power is higher than or quite comparable to many other breast cancer risk factors reported in the literature and/or used in many epidemiology studies based on breast cancer risk prediction models. In addition, since commercialized CAD schemes of mammograms are currently available and used in the clinical practice, computing and utilizing this new imaging marker, if successful in the future validation of the prospective studies using the large and diverse databases, is a cost-effective approach to help improve the efficacy of breast cancer screening using mammography.

Chapter4: Developing A Quantitative Ultrasound Image Feature Analysis Scheme To Assess Tumor Treatment Efficacy Using A Mouse Model

4.1 Introduction

Before performing clinical trials on cancer patients, mouse models are frequently used as an essential step in biomedical research to screen and test new investigative chemotherapy drugs and/or therapeutic methods in order to identify effective drug agents, drug delivery methods, and other treatment technologies for improving the efficacy of cancer treatment [77]. The advantages and necessity of applying mouse models in the initial steps of developing new drugs and/or cancer treatment methods have been extensively investigated and discussed in previous studies [78, 79]. As a result, a large number of mouse models bearing different types of simulated carcinoma tumors have been developed and used in the cancer research field [80-82].

In order to non-invasively visualize and characterize tumor response and/or tissue changes during and/or after cancer treatment, medical imaging plays a vital role by helping validate certain study hypotheses [83]. Many imaging modalities, such as x-ray imaging, including micro-computed tomography (μCT), magnetic resonance imaging (MRI), nuclear and optical imaging, and ultrasound imaging, have been proposed and used for this purpose in recent years [84-86]. Each imaging modality has its specific features and limitations in predicting or assessing tumor response efficacy to the treatment. Compared to other imaging modalities, ultrasound has a number of unique characteristics, making it a more attractive tool to predict or assess cancer prognosis to some clinicians. It is a portable, safe (no harmful radiation), easy-to-use, and low-cost imaging modality to monitor and assess tumor response and tissue characteristics change prior to and post-

treatment [34, 87]. However, despite the potential advantages of using ultrasound imaging to assess treatment efficacy, ultrasound often includes higher noise resulting in a relatively low signal-to-noise ratio. Reliably detecting and computing quantitative image features of tumors from ultrasound images is considered more difficult than computing image features from other imaging modalities (i.e., μCT and MRI). As a result, the feasibility of developing or identifying new quantitative imaging markers computed from ultrasound to predict or assess cancer treatment efficacy at an early stage has not been investigated and validated to date.

Thus, based on the concept and scientific premise of Radiomics [88], the objective of this study is to test the feasibility of identifying and extracting new quantitative image features or markers computed from ultrasound images to predict the efficacy of cancer treatment at an early stage. In order to achieve the study objective, I developed an interactive computer-aided detection (CAD) scheme with an easy-to-use graphic user interface (GUI) to process ultrasound images acquired from the colon carcinoma tumor bearing mice treating with a variety of thermal therapies. The CAD scheme computes a large pool of image features based on tumor morphology, density distribution, and texture-related features from the segmented tumor regions depicted on the ultrasound images. Data analysis was then performed to identify top image features and their fusion method to generate new quantitative imaging markers to predict and compare the efficacy of the thermal therapies under tests at an early stage.

4.2 Materials and Methods

4.2.1 Image Dataset

In this study, I assembled an experimental dataset that includes 23 athymic nude mice bearing C26 adenocarcinomas. These mice were treated with seven different thermal based therapies that combine the focused ultrasound-induced mild hyperthermia and chemotherapeutic nanoparticle formulations. Specifically, these seven treatment methods include (1) high intensity focused ultrasound (HIFU), (2) the chemotherapeutic drug doxorubicin (DOX), (3) DOX and HIFU, (4) low-temperature sensitive liposomes (LTSL), (5) HIFU and LTSL, (6) echogenic low-temperature sensitive liposomes (E-LTSL), and (7) HIFU and E-LTSL. Both LTSL and E-LTSL are different nanoparticle formulations that encapsulate DOX. The details of these treatment methods have been previously reported [89].

In brief, HIFU is a non-invasive therapeutic technique that uses focused ultrasound energy to heat a targeted region of tissue in a controlled manner. DOX is a commonly administered, clinically available chemotherapeutic drug, which is often used for treating a wide range of different cancer types. The addition of HIFU hyperthermia increases blood flow to the tumor region and increases drug perfusion in a targeted manner. LTSLs are thermosensitive liposomes that carry a DOX payload to the tumor site; should heating be applied via HIFU (39-42ºC), DOX will be released at the tumor site, granting the DOX an additional degree of targeting effect. E-LTSLs have the same lipid composition as the aforementioned LTSL and incorporate an ultrasound contrast agent that becomes echogenic (visible on ultrasound) during HIFU heating. They also have the added advantage of promoting improved drug penetration via HIFU/nanobubble interaction.

In the mouse model used in this study, all animal-related procedures were approved and carried out under the guidelines and regulations of the Oklahoma State University Animal Care and Use Committee (ACUP VM-13-24). Specifically, in order to establish a mouse model of colon cancer, C26 cells were grown as a monolayer to 80–90% confluence in RPMI supplemented with 10% v/v fetal bovine serum (FBS) and 1% v/v streptomycin/penicillin. Confluent cells were harvested, washed, and diluted with sterile, cold PBS to generate 0.5×10^5 cells/50 μl. Next, 50 μl of cell inoculum was injected per mouse in the thigh region of the mouse hind leg using a 25-gauge needle (BD, Franklin Lakes, NJ, USA). Mice were monitored, and tumor growth was measured by serial caliper measurements (General Tools Fraction+™, New York, NY, USA). Tumor volumes were calculated using the equation of (length X width^2)/2, where length is the largest dimension and width is the smallest dimension perpendicular to the length. After three days, tumors typically grow and reach a treatment volume of greater than 50mm^3 [89].

In this study, each mouse was treated twice using the targeted thermal therapy on Day 3 and Day 6 after the cell inoculum was injected, respectively. The longitudinal ultrasound images were taken prior to and post-treatment in these two days using a Vevo 2100 ultrasound imaging system at a frequency of 21 MHz [89]. During the process, mice were anesthetized and held in custom-built mouse holders attached to a 3D positioning stage, and the tumors were positioned so that the target was in the center of the focal zone of the ultrasound imaging transducer. In each image acquisition process, multiple ultrasound image frames or a series of imaging videos were taken and recorded. Each mouse was monitored for ten days. On Day 10, the mouse was sacrificed. The tumor was resected, and the tumor size was measured.

4.2.2 Developing CAD Scheme

In this study, I developed an interactive computer-aided detection (CAD) scheme with a graphic user interface (GUI) platform. Following three steps were taken to perform image processing and feature computation. First, after uploading a complete set of ultrasound imaging video series into the GUI, the operator (i.e., a research assistant in the study laboratory) selects one ultrasound image (representing the best frame in the video) in which the tumor area is considered clearly visible. The tumor region is then segmented manually from the ultrasound image. **Figure 4-1** illustrates an example of the tumor regions and their boundary contours segmented from 4 sets of ultrasound images acquired from one mouse prior and post-DOX treatment on Day 3 and Day 6 using the GUI of our CAD scheme, respectively. Specifically, I used the algorithm illustrated in **Figure 4-2** for processing the images.

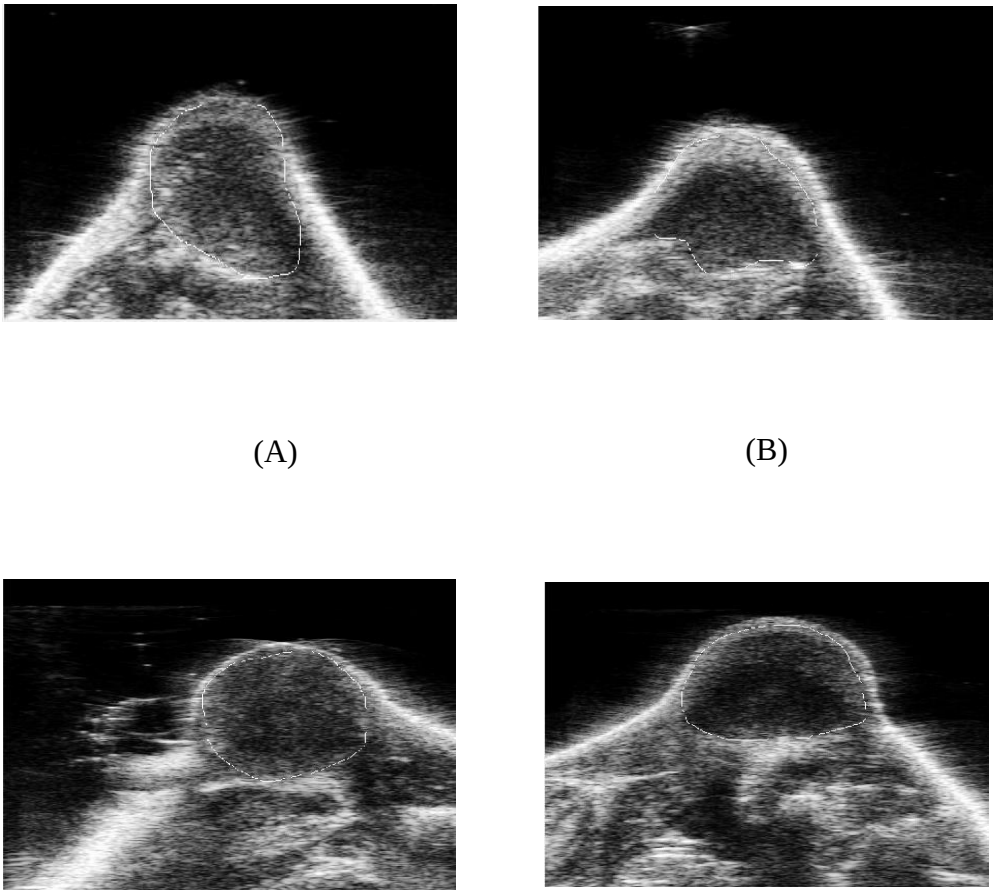


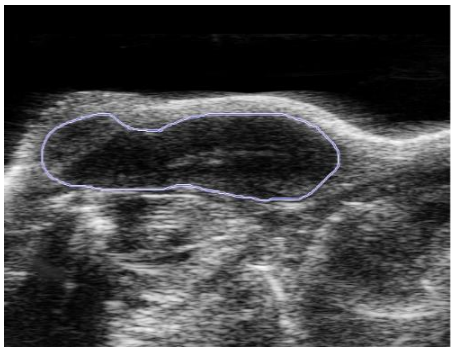
Figure 4-1. An example of four ultrasound images taken from a mouse in Day 3 (A) prior- DOX treatment and (B) post-treatment, in Day 6 (C) prior-treatment and (D) post-treatment, respectively. The tumor boundary contours are marked on each image.

```
I=Total frames
 For (i=1: number of total frames)
      J=Selected frame by the operator
      J´=Manual segmentation on J
 end
```

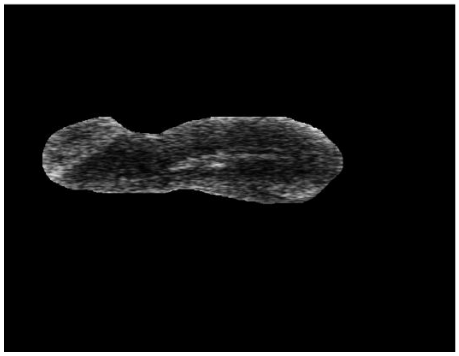
Figure 4-2. Proposed Algorithm for processing each image.

4.2.3 Feature Extraction

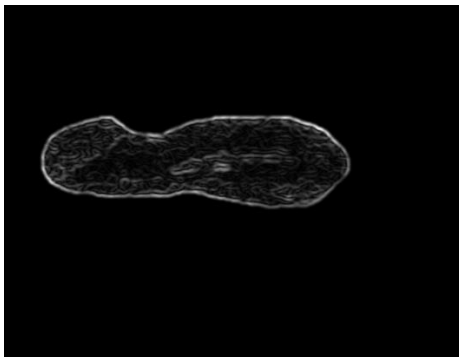
After tumor segmentation, the CAD scheme applies a low pass Gaussian filter on images to decrease the ultrasound images' inherent noise (**Figure 4-3**). CAD then computes image features. A total of 284 image features are computed from each segmented tumor region. Similar tumor-related image features have been computed from other imaging modalities (i.e., CT and MRI) in our previous studies to develop quantitative image markers for predicting tumor response to chemotherapies of treating breast and ovarian cancer [70, 90]. These features can be categorized into four groups as summarized in **Table 4-1**, which include (1) nine morphology-based image features; (2) 21 tumor density distribution related image features; (3) 44 grayscale run length (GSRL) based texture related image features [91], which include (a) Short Run Emphasis (SRE), (b) Long Run Emphasis (LRE), (c) Gray-Level Nonuniformity (GLN), (d) Run Length Nonuniformity (RLN), (e) Run Percentage (RP), (f) Low Gray-Level Run Emphasis (LGRE), (g) High Gray-Level Run Emphasis (HGRE), (h) Short Run Low Gray-Level Emphasis (SRLGE), (i) Short Run High Gray-Level Emphasis (SRHGE), (j) Long Run Low Gray-Level Emphasis (LRLGE), and (k) Long Run High Gray-Level Emphasis (LRHGE) computed in four different directions (0°, 45°, 90°, and 135°), respectively; and (4) 210 image features computed from the wavelet transformation maps. Appendix I illustrated features' computations.



(A)



(B)



(C)

Figure 4-3. Illustration of applying Gaussian filter to the ultrasound image, which shows (A) manually marked tumor boundary contour, (B) the segmented tumor region, and (C) tumor image after applying the Gaussian filter.

Table 4-1. List of the computed 284 image features in four feature groups.

Feature Class	Feature Number	Feature Description
Morphology	1-9	Volume, convexity, maximum radius, radius standard deviation (STD), surface area, compactness, maximum three-dimensional diameter, spherical disproportion, and spherical ratio.
Density	10-30	Density, density STD, gradient mean, gradient STD, ISO-intensity, fluctuation mean, fluctuation STD, mean contrast, contrast, skewness, kurtosis, STD ratio of tumor to the boundary, energy, entropy, maximum intensity, mean absolute deviation, median, minimum, range, RMS, and uniformity.
Texture	31-74	11 gray-level run length-based features in four directions (0°, 45°, 90°, and 135°).
Wavelet	75-284	Apply the density and texture features to the four wavelet decompositions.

Precisely, to compute image features in group four, CAD applies the wavelet transform on the ultrasound image so that the image is decomposed into four components, including I_{LL}, I_{LH}, I_{HL}, and I_{HH}, where H and L are labeled as the high- or low-scale decomposition in either the X or Y direction. Intrinsically, in this computation, I_{HL} denotes the component after applying the high-scale and low-scale filter along with the X and Y directions, respectively. For each component, the

density and texture features measured in the second and third group are recalculated, respectively. **Figure 4-4** shows a two-step algorithm to filter images and compute image features.

After image processing and feature computation, I assembled four initial image feature pools, which include the image features computed from (1) prior treatment on Day 3, (2) the difference between prior and post-treatment on Day 3, (3) prior treatment on Day 6, and (4) the difference between prior and post-treatment of Day 6. Afterward, all computed features in each feature pool were normalized to the values between 0 and 1. Hence, two initial feature pools were established for each thermal treatment on Day 3 or Day 6. The first one includes image features extracted from prior treatment ultrasound images only, and the second one includes image feature difference by subtraction between two features computed from the two matched images acquired prior and post-thermal therapy of the same mouse.

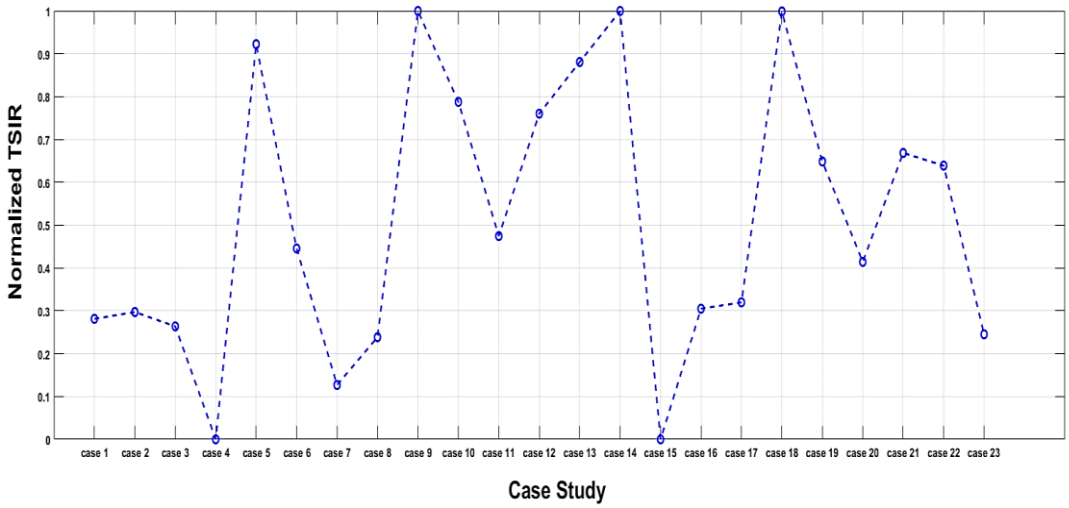
```
J´´=Low pass Gaussian filter (J´)
Feature_vector = CAD(J´)
```

Figure 4-4. The proposed algorithm for image filtering and feature computation.

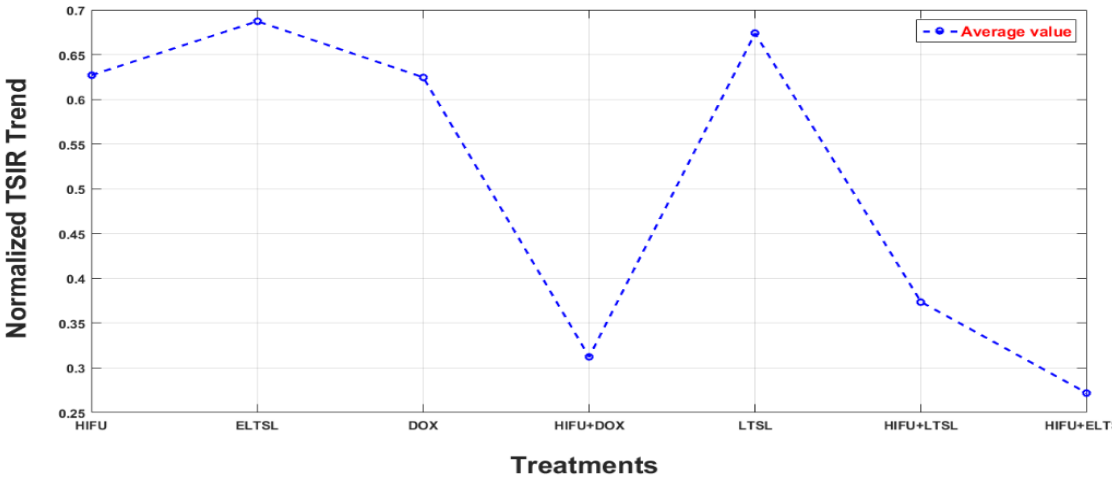
In order to identify and select the potentially effective quantitative image features or markers, I used tumor size change during the period of starting tumor treatment (Day 3) to the end of monitoring (Day 10) as a comparison reference ("ground-truth" of treatment efficacy) in this study. Specifically, the tumor size increment ratio (TSIR) for each mouse is computed with the **Formula 4-1** during this period.

$$TSIR = \frac{x2-x1}{x2} * 100 \qquad (4\text{-}1)$$

where x1 and x2 are the mouse tumor size at day 3 and day 10, respectively. This TSIR based evaluation criterion is similar to Response Evaluation Criteria in Solid Tumors (RECIST) guidelines used in current clinical practice to assess tumor response to therapies among cancer patients [92]. Then, the computed TSIRs were normalized between 0 and 1. **Figure 4-5** shows the normalized TSIR for all 23 mice and is averaged among all therapy groups that were used in this study.



(A)



(B)

Figure 4-5. Distribution of the normalized TSIR ratios based on (A) each mouse and (B) average of each therapy group.

In order to identify the association between the image features and treatment efficacy, I computed the correlation coefficient of each feature with TSIR using the following equation [80]:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \qquad (4\text{-}2)$$

where “r” is the Pearson correlation coefficient, “x” and “y” are one selected image feature and the TSIR, respectively. “n” is the case study size, which is 23 in this study. Interpretation of the computed Pearson’s Correlation coefficient to the association between two compared parameters (i.e., one image feature and TSIR in this study) is listed in **Table 4-2** [93]. Thus, the image features that have higher Pearson’s correlation coefficients with TSIR indicate the higher performance to predict treatment efficacy in this study.

Table 4-2. Pearson Correlation coefficient interpretation [55].

“r “ Value	**Relation**
+0.70 or higher (-0.70 or lower)	Very strong positive (negative) relationship
+0.40 to +0.69 (-0.40 or -0.69)	Strong positive (negative) relationship
+0.30 to +0.39 (-0.30 or -0.39)	Moderate positive (negative) relationship
+0.20 to +0.29 (-0.20 or -0.29)	Weak positive (negative) relationship
+0.01 to +0.19 (-0.01 or -0.19)	No or negligible relationship
0	No relationship [zero correlation]

By computing Pearson’s correlation coefficients of 284 image features stored in each of the four initial feature pools representing the ultrasound imaging tests performed on Day 3 and Day 6, I first selected the top five image features that have a higher correlation with TSIR in each feature pool. Next, I calculated the correlation coefficient of these five features with each other. Then, in order to reduce redundancy, I selected two features among these five top features, which have the

lowest correlation coefficient, to generate a new fusion marker using an equally weighting method, $F_{new} = (F_1 + F_2)/2$. A similar fusion method has been used in our previous studies (i.e., [94]).

I also recognize that unlike other imaging modalities (i.e., CT or MRI), an ultrasound imaging test usually acquires multiple image frames. In this study, one ultrasound imaging test or scan typically includes up to 200 image frames. In order to test the scientific rigor or reproducibility of the quantitative image features computed from different ultrasound image frames, I defined one frame initially selected by the operator of the GUI of our CAD scheme as the base frame of a set of ultrasound images acquired in one test. I also processed and computed the same image features of the segmented tumor region from all other ultrasound image frames (i.e., the remaining 199 frames) in this set. Then, I computed the mean correlation coefficient and the standard deviation of the features computed from the base frame and all other 199 frames. **Figure 4-6** shows an algorithm to examine the reproducibility of the features computed from the base frame as compared to the features computed from all other frames.

```
K=Total frames
N=Manual segmentation on frames
base_feature_vector=CAD(N)
    For=j: total number of frames
        N´=map segmentation of N on frame (j)
        N´´=correction(N´)
    end
```

Figure 4-6. The proposed algorithm for examining reproducibility or consistency between the image features computed from the base frame and other frames.

4.3 Results

Table 4-3 shows two sets of five optimal image features with the highest Pearson correlation coefficients with the treatment outcome (TSIR) and the corresponding p-values. These features are selected from the two initial image feature pools that record the image features computed from the ultrasound images acquired prior treatment on Day 3 and Day 6, respectively. It shows that using image features computed from prior treatment ultrasound images acquired on Day 3 yielded a moderate correlation, and there are no significant differences between the top five features ($p > 0.05$) while using the image features computed from prior treatment ultrasound images acquired on Day 6 increase the correlation level to a strong positive correlation and the number one feature yielded significantly higher correlation as compared to other four top features ($p < 0.01$).

The top five performed image features selected on Day 3 and Day 6 are different, as shown in **Table 4-3**, which indicates that treatments have an impact on the change of tumor morphological and texture characteristics. In addition, **Table 4-4** shows and compares five sets of correlation coefficients of the same image features computed from prior treatment ultrasound images acquired on both Day 3 and Day 6.

The results show that image features contain increased discriminatory power or higher correlation coefficients as they approach the endpoint of Day 10 (i.e., Day 6 vs. Day 3) to predict treatment efficacy or outcome.

Table 4-3. List of two sets of the selected 5 top image features from 2 image features of prior treatment on Day 3 and Day 6.

Day 3			Day 6		
Feature	Correlation coefficient with TSIR	P-value comparing to Range	Feature	Correlation coefficient with TSIR	P-value comparing to GLN_{HL}
Range	**0.375**		GLN_{HL}	**0.680**	
$Entropy_{HH}$	0.361	0.468	RP_{HL}	0.643	<0.01
RP_{LL}	0.359	0.415	GLN_{HL} 90°	0.605	<0.01
$Entropy_{LL}$	0.355	0.478	RLN_{LL}	0.598	<0.01
GLN_{HL} 90°	0.344	0.377	$Entropy_{HL}$	0.597	<0.01

Table 4-4. Comparison of the correlation coefficients of the same image features computed from prior treatment ultrasound images acquired on Day 3 and Day 6.

	GLN_{HL} 90°	**Tumor Vol.**	**RLN_{HL}**	**RP_{HH} 90°**	**GLN_{LL}**
Day3	0.344	0.341	0.329	0.326	0.318
Day6	0.605	0.525	0.586	0.546	0.551

Table 4-5 shows the correlation coefficients between TSIR and five top image features selected from the feature pool that contains image feature differences computed between prior and post-treatment ultrasound images acquired on Day 3. It shows that using the image features that represent the difference in tumor response or characteristic change prior and post-therapies yields

a higher correlation with TSIR or higher prediction power. In addition, by selecting two of the five top features (as listed in **Table 4-5**), which have the smallest correlation coefficients among the five top features, I applied an equally weighted fusion method to generate a new image marker. The correlation coefficient of this new fusion-based image marker and TSIR significantly increased to 0.679, as shown in **Table 4-5**.

Table 4-6 shows the five selected image features computed from the differences between the prior and post-treatment ultrasound images acquired on Day 6 with the highest correlation coefficients to TSIR. However, when comparing the correlation coefficients of the top five-

Table 4-5. List of the five selected image features computed from the difference of prior and post-treatment ultrasound images acquired on Day 3 with the high correlation with TSIR.

No.	**Features**	**Correlation coefficient with TSIR**	**P-value comparing to F1**
F1	GLN_{HL}	**0.552**	
F2	LGRE 0°	0.495	0.128
F3	$Range_{HL}$	0.388	0.011
F4	$LGRE_{LL}$	0.387	0.530
F5	Gradient STD_{LH}	0.373	0.809
	Fusion Average (F1& F3)	**0.679**	**<0.01**

features computed from prior treatment ultrasound images acquired on Day 6 (**Table 4-3**), correlation coefficients decrease, which indicates that adding the image features computed from

post-treatment ultrasound images on Day 6 does not help increase power to predict treatment outcome (TSIR determined on Day 10).

Table 4-6. The top image features were computed from the difference of prior and post-treatment ultrasound images acquired on Day 6 with the high correlation with TSIR.

Feature	**Correlation coefficient with TSIR**	**P-value comparing to** RLN_{HH} 45°
RLN_{HH} 45°	**0.428**	
RLN_{HH} 135°	0.420	<0.01
RP_{HH} 90°	0.358	<0.01
RP_{HH}	0.357	<0.01
$LGRE_{LL}$	0.306	0.251

Subsequently, I separately analyzed and sorted feature distribution in seven groups of different thermal therapy methods. Each of the seven treatment groups involves 2 to 4 mice. For example, **Figure 3-7** shows the distribution (or boxplot) of one of the top image features (the GLN_{HL} values) computed from the difference of prior and post-treatment ultrasound images acquired on Day 3 across the treatment groups. The result shows a trend indicating that the image feature values vary when using different thermal therapy methods. Finally, **Table 3-7** shows examples of the mean correlation coefficient, standard deviation, and 95% confidence interval of the base frame and the other 199 frames for the top five image features computed from the difference of prior and post-treatment ultrasound images acquired on Day 3. These image features were computed from a

mouse within the HIFU+ELTSL treatment group. The results show the image features computed from different frames of ultrasound images acquired in one test are highly correlated or invariant.

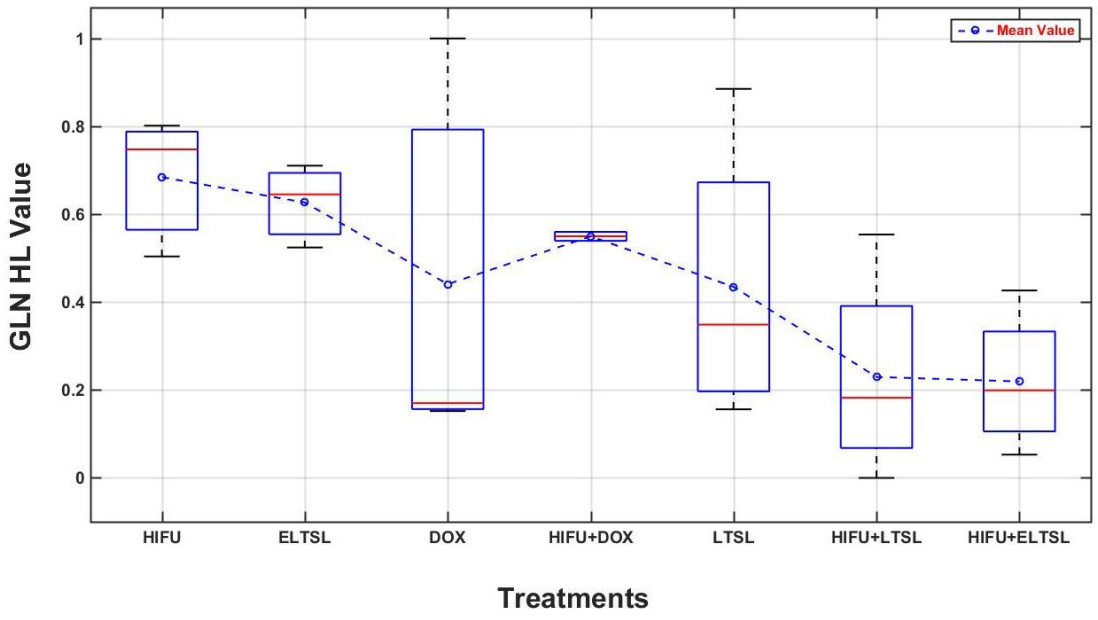


Figure 4-7. The GLN HL values computed from all mice under different treatments, which are sorted from low to high performance (right to left), respectively.

Table 4-7. An example of the base frame and other frames relationships.

Features	Mean Correlation Coefficient	95% Confidence Interval	Standard deviation
GLN_{HL}	0.978	[0.951, 1.00]	0.125
LGRE 0°	0.9668	[0.019, 0.917]	0.224
$Range_{HL}$	0.9491	[-0.058,0.608]	0.167
$LGRE_{LL}$	0.9719	[0.100, 0.978]	0.219
Gradient STD_{LH}	0.9513	[0.089, 0.905]	0.204

4.4 Discussion

In cancer research, many previous studies have reported to develop and apply either molecular biomarkers (i.e., [95-97]) or quantitative image markers (i.e., [46, 69, 70, 90, 98]) to predict tumor response to chemotherapies and/or other therapeutic methods at an early stage. In this study, I investigated and demonstrated the feasibility of identifying new quantitative image markers computed from ultrasound images. This study has a number of unique characteristics and potential impacts. First, in the previous studies, quantitative image markers were computed based on the Radiomics concept that uses CT and/or MRI images to predict cancer prognosis or tumor response to treatment [88]. In this study, I applied the Radiomics approach to ultrasound images to identify new quantitative image markers to predict cancer treatment efficacy. Due to ultrasound imaging's advantages as a diagnostic modality, developing highly performed and robust image markers could be a cost-effective approach in future research and clinical service. Second, thermal-based

therapies have been emerging as a promising cancer treatment method. However, accurate prediction aimed at determining the efficacy or treatment outcomes of different therapeutic approaches remains an unsolved challenge. This is the first study with a computer-aided approach to developing new quantitative imaging markers that can predict thermal therapies' efficacy using a mouse model. Third, I computed and compared image features of the longitudinal images acquired from prior and post-treatment ultrasound images and two-time points (Day 3 and Day 6 after initial tumor cell embedment). As a result, I am able to conduct a more comprehensive data analysis and identify an optimal approach to extract and compute image markers in order to more accurately predict treatment efficacy or outcome at an early stage.

From the experiments and data analysis results, I can make the following new observations. First, this study shows that it is possible to identify quantitative ultrasound image feature based markers at an early stage (i.e., Day 3 in this study) to predict thermal therapy efficacy. However, in the early stage, using the image marker computed from both prior and post-treatment ultrasound images can yield substantially higher prediction accuracy as compared to using the prior treatment images only (i.e., correlation coefficients of 0.375 vs. 0.679 as shown in **Tables 3-3** and **3-5**). This result is consistent with our earlier study of developing quantitative image markers computed from prior and post-chemotherapy CT images to predict ovarian cancer patients' response to chemotherapy in the clinical trials [70].

Second, I observed that using quantitative image markers computed from the prior treatment ultrasound images acquired on Day 6 yielded substantially higher prediction power than the image markers computed from the prior treatment ultrasound images acquired on Day 3. The correlation coefficients increase from the moderate level to the strong positive level for both top five features, as shown in Table 3-3, and the same highly performed features shown in **Table 3-4**. Thus, the

trend is consistent, which indicates that the images acquired at the later stage (i.e., Day 6 in this study) contain higher discriminative information or predictive power to evaluate treatment outcomes.

Third, I also observed a different phenomenon when using the image features computed from the difference between prior and post-treatment ultrasound images acquired on Day 3 and Day 6. Unlike the image markers computed on Day 3, adding post-treatment ultrasound images acquired on Day 6 does not further increase prediction power or have lower correlation coefficients, as shown in **Table 3-6**. It reveals that in the earlier day (Day 3), the tumors have higher positive responses to the thermal therapies than those on a later day (Day 6), which may indicate that at the later stage, tumors are more resistant to the treatment. Thus, the observation may clearly show that early treatment (i.e., on Day 3) is more important and effective than later treatment (i.e., on Day 6 in this study), which is consistent with the established scientific evidence in cancer treatment research.

Fourth, the computed image feature values also vary on the ultrasound images acquired from the mice under different treatment methods. In this study, 23 mice were treated with seven different thermal therapies. Based on the final tumor size measurement results or TSIR on Day 10, the effectiveness of these seven thermal methods has a monotonically decreased trend from thermal therapy method 1 (HIFU) to method seven (HIFU+E-LTSL). The computed image features also show a similar trend, as shown in **Figure 3-7**, which indicates the high correlation between the image features and thermal therapy methods. Thus, using quantitative image markers also has the potential to help identify optimal therapy methods.

Last, although computer-aided detection schemes of medical images can be quite sensitive to change of image noise [99], I observed that image features computed from all image frames (i.e.,

200 in this study) acquired at one ultrasound imaging test of a mouse were highly correlated or invariant to the small change of inherent image noise. As an example shown in **Table 3-7**, the mean correlation coefficients of the top five image features computed from the base image frame and other 199 image frames in one ultrasound imaging scan of a mouse ranged from 0.949 to 0.978. The results revealed that as long as using a well-established or controlled imaging protocol, it is feasible to robustly compute quantitative image features from the ultrasound images for predicting the efficacy of tumor response to therapies.

4.5 Conclusion

Ultrasound is a safe, easy-to-use, and low-cost medical imaging modality. In this study, I investigated the feasibility of identifying and applying quantitative image feature based markers computed from ultrasound images using a C26 adenocarcinoma bearing mouse model to assess tumor response to thermal therapies. The study results demonstrated that ultrasound images acquired prior and post-therapy at an early stage (i.e., Day 3 of this study) contained useful and highly discriminative information that can possibly predict tumor response to therapies at an early stage. Although ultrasound images may have a higher inherent noise level than other medical imaging modalities (i.e., CT or MRI), this study also indicated the possibility of computing highly robust image features to develop robust image feature based markers. Thus, based on the preliminary results of this study, future studies using larger and more diverse image datasets are needed to further validate the performance and potential utility of this methodology and similar approaches.

Chapter 5: Applying A Novel Feature Selection Method To Optimize The Machine Learning Model For Peritoneal Metastasis Prediction In Gastric Cancer Patients

5.1 Introduction

While the incidence of gastric cancer has reduced recently, it remains the third leading cause of cancer-related death worldwide [100]. Despite the fact surgery remains the only therapeutic treatment option, preoperative neoadjuvant chemotherapy (NAC) has proven promising results with increased curative resection rates and improved survival [101]. In order to prevent the unfavorable effect of NAC, patients with different disease stages must be distinguished from each other [102] because, for each step of the disease, the treatment would be different [103]. Recent studies revealed that applying preoperative NAC for advanced gastric cancer patients with peritoneal metastasis (PM) yielded a much better clinical outcome and lead to a more significant overall survival rate [104, 105]. Therefore, an accurate assessment of the PM's existence is vital for selecting proper patients for undergoing NAC. Considering the overall accuracies of subjectively reading computed tomography (CT) and endoscopic ultrasound images are not entirely reliable [102, 103], an alternative technique is required to facilitate the assessment of tumor stages and the risk of PM.

Recently, the novel radiomics technique has been applied in extracting quantitative information from medical images with a high dimensional image feature. Consequently, the data mining of the image feature pool suggests an efficient method to develop machine learning (ML) models and predict clinical outcomes [39, 88]. While numerous radiomics based ML models have

been reported to differentiate and stage gastric cancer patients [106, 107], these studies extracted radiomics features from the tumor region that is segmented manually from one CT slice chosen by the radiologist. Hence, the correlation analysis based method was used to determine a small set of image markers, which cannot eliminate the redundant features. Therefore, the discriminatory power and prediction accuracy of these ML models were inadequate. To overcome such restrictions, in this study, I propose to develop a new computer-aided detection (CAD) scheme in order to predict the risk of PM among gastric cancer patients. First, our scheme segments automatically gastric tumor volume in 3D CT image data, leading to better calculating image features related to the heterogeneity of the gastric tumors. Second, to decrease feature space dimensionality and better ascertain non-redundant image features from a large initial pool of extracted radiomics features, I examine and apply a random projection algorithm (RPA). Third, to prevent bias in feature vector generation, RPA is embedded in a multi-feature fusion-based machine learning (ML) model for prediction of the PM risk, which is trained and tested using (1) a synthetic minority oversampling technique (SMOTE) and (2) a leave-one-case-out (LOCO) cross-validation method. The details of the study, including experimental measures, data analysis results, and discussions, are presented in the following sections of this chapter.

5.2 Materials and Methods

5.2.1 Image Dataset

In this study, I retrospectively use a dataset of abdominal computed tomography (CT) images. To prevent potential case selection bias, the dataset initially includes 219 consecutive patients diagnosed and treated with gastric cancer. Afterward, by eliminating the unresectable or undetectable cases based on CT examinations and poor image quality as verified by the

radiologists in the retrospective review, 159 cases are involved in this study dataset. Among these cases, 121 patients have PM, and 38 patients do not have PM. **Table 5-1** summarizes the distribution of overall demographic information and related clinical results of these 159 gastric cancer patients in this dataset.

Table 5-1. The overall demographic information and related clinical results of patients in the dataset.

	Category	Cases with PM	Cases without PM
Total Cases		121	38
Age (years old)	< 45	11 (6.9%)	5 (3.1%)
	45 – 65	72 (45.2%)	23 (14.4%)
	> 65	38 (23.8%)	10 (6.2%)
	Mean ± SD	59.49±11.97	59.11±8.75
	Median	61	60
Gender	Men	94 (59.1%)	
	Women	27 (16.9%)	

			8 (5.0%)
Tumor Location	Upper	37 (23.2%)	19 (11.9%)
	Medium	20 (12.6%)	7 (4.4%)
	Lower	50 (31.4%)	12 (7.5%)
	Diffuse	14 (8.8%)	0
Pathological Staging after Surgery	I	0	38 (23.9%)
	II	26 (16.4%)	0
	III	32 (20.1%)	0
	IV	63 (39.6%)	0
Bormann Type	1	1 (0.6%)	0
	2	21 (13.2%)	11 (6.9%)
	3	94 (59.1%)	25 (15.7%)
	4	5 (3.1%)	2 (1.3%)

Each patient has undergone an abdominal CT imaging examination during the original cancer diagnosis before surgery. All CT examinations were carried out using a multidetector CT machine (GE Discovery CT750 HD; GE Revolution CT, GE Healthcare). Each patient is asked to be fast from food overnight and drank 600-1000 ml water to distend the stomach before the CT examination. The contrast-enhanced CT images are taken a delay of 28 s (arterial phase), 55 s (portal phase), and 120 s (venous phase) after administration of infused 1.5 ml/kg body weight iodinated contrast agent (Optiray 320 mg I/mL, Bayer Schering Pharma) intravenously at a flow

rate of 2.5 ml/s. The CT scanning parameters include (1) tube voltage switching between 120 kVp and 140KVp in spectral imaging mode, (2) tube current automatically augmenting with the maximum limit of 200mA, (3) tube rotation time of 0.76 – 0.80s, (4) detector collimation of 64 × 0.625 mm, (5) field of view with 350 – 500 mm, and (6) the image matrix with 512 × 512 pixels and reconstruction thickness of 2.5 mm.

5.2.2 Tumor Segmentation

By identifying the heterogeneity of tumors in the clinical images, I improved and applied a hybrid tumor segmentation scheme that utilized a dynamic programming method [108, 109] to identify growing thresholds of a multi-layer topographic region growing algorithm adaptively and primary contour in active contour algorithm. In particular, the tumor segmentation scheme contains the following steps; First, a Weiner filter is used to lessen image noise. Second, a primary seed is placed at the center of the tumor region of a CT slice in which the tumor has its largest area. To decrease inter-operator variability in selecting the initial seed and improve the robustness of segmentation results explained in the previous study [110], a predefined window with the size of (5,5) around the initial seed is automatically generated. A pixel with the minimum value within the window is identified and selected as the initial seed point. Third, to automatically define the first threshold value for the region growing algorithm, a new predefined window with size of (5,5), which confirms to fully locate inside all tumor regions of our dataset and prevent potential risk of growing leakage at the first growth layer, is generated around the new seed point. Subsequently, the scheme calculates the pixel value differences between the center pixel and boundary pixels and detects the maximum difference. Consequently, the region growing threshold is defined as $T_1 = V_c + 0.25 \times D_{max}$, where V_c is the central pixel value and D_{max} is the calculated maximum pixel

value difference inside the bounding window. This threshold value is applied to determine the first layer of region growing to segment the tumor region, depicting on one CT image slice.

Fourth, after defining the first layer of tumor region growth, the growing threshold of the second layer is computed by $T_2 = T_1 + \beta C_1$ where C_1 is the constant value of the first layer, and β is a coefficient (i.e., 0.5). This multi-layer region growing algorithm continues to grow until the growing ratio between two adjoining layers is bigger than two times the last layer's size. Finally, after the region growing algorithm stops, the scheme chooses the boundary contour of the last region growing layer as the initial region contour. The active contour algorithm is followed to expand or shrink the contour curve for the best fitting tumor boundary.

Figure 5-1 and **Figure 5-2** show the block diagram of this tumor segmentation scheme and an image example of employing the above phases to segment a tumor region depicting on one CT slice, respectively.

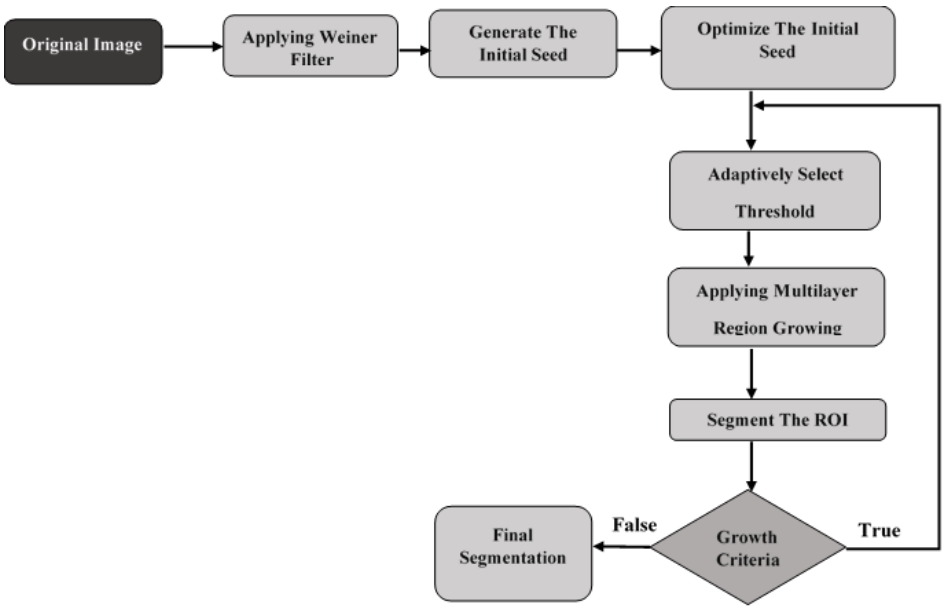
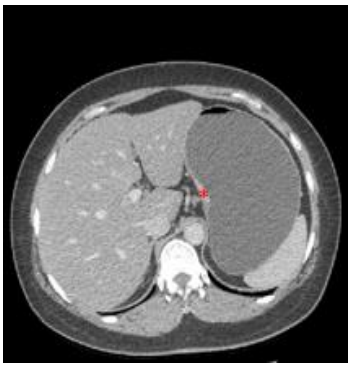
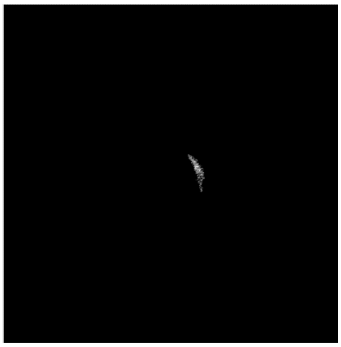


Figure 5-1. The block diagram of the 2D tumor segmentation.

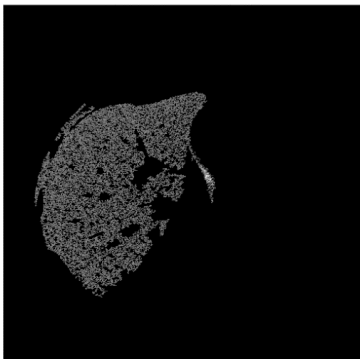


(a) Selecting the initial seed

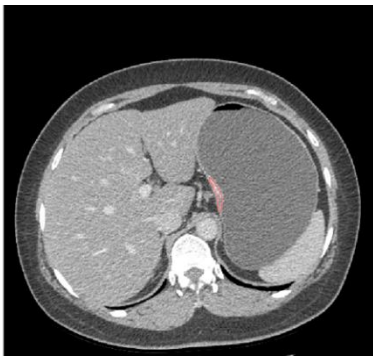
(b) Applying region growing based on auto threshold



(c) Continuing to grow till meeting the growth criteria



(d) Growing until the criteria of the growth do not meet



(e) Final segmentation

Figure 5-2. The process of 2D tumor segmentation.

Subsequently, after tumor region segmentation on one CT slice, the CAD scheme continues to apply tumor region segmentation by scanning in both up and down directions until no tumor region is found in the next adjacent CT slice. In this progression, the tumor region's central point detected in the adjacent CT slice is mapped into the new CT slice as the initial region growing seed. After that, the tumor region segmentation in this targeted slice is automatically performed from the mapped growing seed. Moreover, a tumor growing boundary condition is restricted by the adjacent slice to facilitate the multi-layer region-growing process and prevent growth leakage.

Figure 5-3 displays an example of tumor regions' segmentation showing several CT image slices of one case. In this way, 3D tumor volume can be segmented and computed.

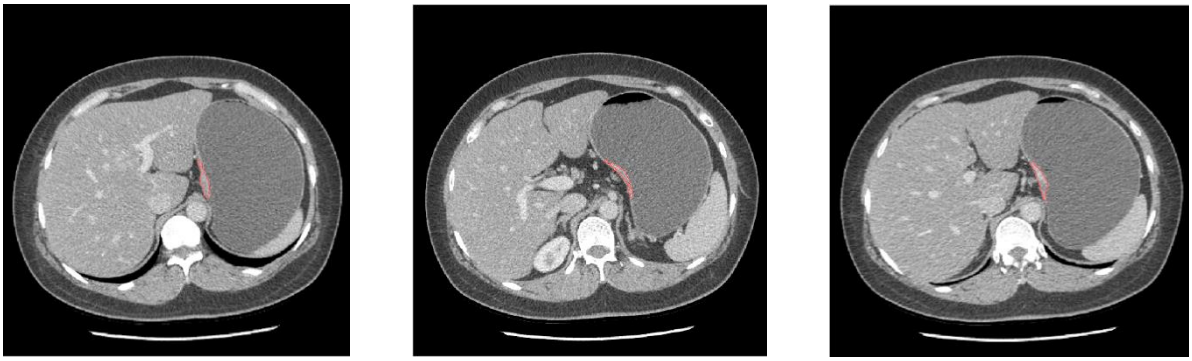


Figure 5-3. An Example of 3D segmentation of a tumor region in 3 different slices.

5.2.3 Feature Extraction

Once 3D tumor volume is segmented, the CAD scheme is applied to compute a large pool of radiomics-based image markers, including 315 features extracted from each segmented 2D tumor region (ROI) illustrating on one CT image slice. These features were considered into four main groups, including (a) the grayscale-run length (GLRLM) features in which 44 two-dimensional

features are computed. (b) The Gray Level Difference Methods (GLDM) probability density function (PDF) features in which four features of mean, median, standard deviation, and variance from each probability density function representing statistical texture features of ROI are computed. (c) Wavelet domain features in which the image is first decomposed into four main components consist of low and high scale decomposition by wavelet transform [111]. Then, the GLCM features [112], as well as 21 tumor density [113] and GLDM features [114], are extracted from those components. (d) Laplacian of Gaussian (LoG) features in which a Gaussian smoothing filter is first applied to lessen the sensitivity to the noise, and next, the Laplacian filter sharpens the image's edge and highlights rapid intensity changes inside the region [115]. Following that from the extracted points after applying the LoG filters, the mean, median, and standard deviation are calculated. **Figure 5-4** demonstrates the flow diagram of the feature extraction process.

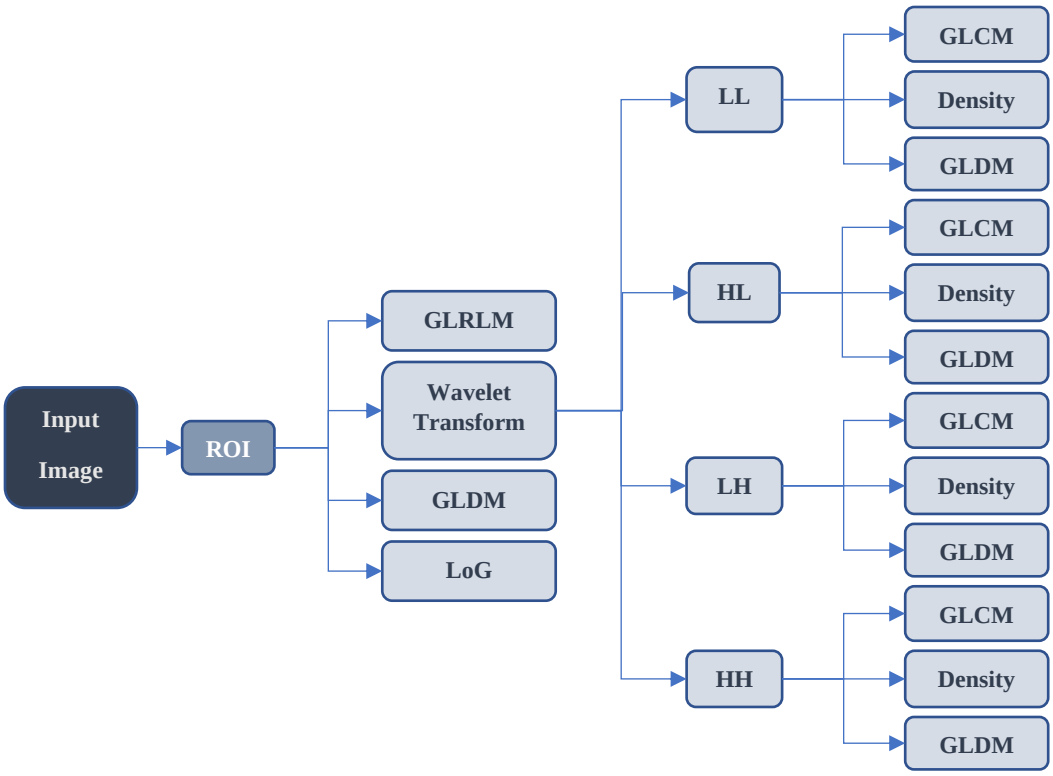


Figure 5-4. Flow Diagram of Feature extraction Method.

Following the 2D features' computation of all segmented tumor regions in N involved CT image slice, CAD scheme computes each 3D feature (F_{3D}^{k}) as:

$$F_{3D}^{k} = \sum_{i=1}^{N} m_i \times F_{2D}^{k} \quad (5\text{-}1)$$

In which m_i is the ratio of the tumor volume segmented on a ith CT slice to the whole segmented tumor volume on all N involved CT slices. The tumor volume segmented on a i th slice is computed by multiplying the segmented region size (2D) to the CT slice thickness. Lastly, all 315 computed 3D feature values are normalized between 0 to 1 to decrease case-based reliance, and weight all features evenly.

5.2.4 Feature Reduction Using Random Projection Algorithm

Since the initial feature pool consists of 315 image features, many of them can be irrelevant (with lower performance) or redundant (highly correlated). Therefore, choosing a small set of optimal features to reduce the feature dimension and increase learning accuracy is essential. This study explores and applies a novel image feature regeneration method of the Random Projection Algorithm (RPA). Theoretic analysis has shown that the RPA has advantages for its simplicity, high performance, and robustness in comparison with other feature reduction methods; however, empirical results are sparse [50]. Meanwhile, RPA has been investigated and tested in many engineering applications such as text [116] and face and object recognition [117] and yielded comparable results to conventional feature regeneration methods like principal component analysis (PCA) [118]. Nonetheless, the advantage of applying RP methods over their alternative is that they generate more robust results and computationally inexpensive [50, 119].

This study will employ RPA to generate optimal image features from the original large set of radiomics features. Following is a brief introduction of the RPA method. By taking into account each case as a point in a *k* dimensional space, where *k* represents the number of image features, the Euclidian distance between two points can be defined as follows:

$$|M - Z| = \sqrt{\sum_{i=1}^{k}(m_i - z)^2} \qquad (5\text{-}2)$$

Considering Formula (2), $M = (m_1, \dots, m_k)$, and $Z = (z_1, \dots z_k)$ are two points in the k dimensional space. Similarly, the volume of a sphere with radius r and volume of V in k dimensional space is expressed as follows in **Formula 5-3** [120]:

$$V(k) = \frac{r^k \pi^{\frac{k}{2}}}{\frac{k}{2}\Gamma(\frac{k}{2})} \quad (5\text{-}3)$$

The normalization of the feature matrix between [0, 1] proposes that all data can be involved in a sphere with a radius of 1. The vital fact about a sphere with a unit radius is that the more enhancement in dimension, the more reduction in the volume (**Formula 5-4**). Concurrently, the possible distance between the two points remains at 2 [120].

$$\lim_{k\to\infty}\left(\frac{\pi^{\frac{k}{2}}}{\frac{k}{2}\Gamma\left(\frac{k}{2}\right)}\right) \cong 0 \quad (5\text{-}4)$$

Furthermore, regarding the theory of the heavy-tailed distribution, for a case like $M = (m_1, \dots, m_k)$ in the space of features, considering features independent with a satisfactory approximation, or roughly perpendicular variables mapping to different axes, with $E(m_i) = p_i$, $\sum_{i=1}^{k} p_i = \mu$ and $E|(m_i - p_i)^d| \leq p_i$ for $d = 2,3, \dots, \lfloor t^2/6\mu \rfloor$, then, a probability can be calculated using **Formula 5-5** [120]:

$$prob\left(\left|\sum_{i=1}^{k} m_i - \mu\right| \geq t\right) \leq Max\left(3e^{\frac{-t^2}{12\mu}}, 4 \times 2^{\frac{-t}{e}}\right) \quad (5\text{-}5)$$

The more the value of t surges, the less chance of a point be out of that distance. Hence, M should be focused on the mean value. Specifically, according to **Formula 5-4** and **Formula 5-5**, with a reasonable estimation, all data are contained in a sphere of unit size, and they are focused around their mean value. As a result, if the dimension increases, the volume of the sphere would close to zero. Consequently, the difference between the cases is not enough for precise classification.

According to the aforementioned analysis, the larger the initial feature vector size, the bigger the space dimension is. Hence, most of the data is focused around the center, which leads to less

difference between the features. Consequently, to reduce the feature dimension, an efficient technique is the one that reduces the features' dimensionality while preserves the distance between the points, representing rough preservation of the vast amount of information. Suppose we apply a conventional feature selection method and accidentally select a d-dimensional sup-space of the initial feature vector. In that case, it is expected that all the projected distances in the new space are within a determined scale-factor of the initial *k*-dimensional space [121]. Accordingly, it is feasible that after removing the redundant features, the accuracy would not improve due to the fact that the divergence between the points is not significant enough to consider as a robust model.

To address the aforementioned concern and improve the feature space, Johnson-Lindenstrauss Lemma's theory can be applied in RPA [122]. This theory expresses that for any $0 < \epsilon < 1$, and for any number of cases as t, which are like the points in k-dimensional space (R^k), if assuming d as a positive integer, **Formula 5-6** can be used to compute this integer number [122]:

$$d \geq 4\frac{ln\, t}{\left(\frac{\epsilon^2}{2}-\frac{\epsilon^3}{3}\right)} \qquad (5\text{-}6)$$

Afterward, for any set W of t points in R^k, for all $z, w \in W$, it is shown that there is a map, or random projection function like $f: R^k \rightarrow R^d$, which keeps the distance defined by **Formula 5-7** [122]:

$$(1-\epsilon)|z-w|^2 \leq |f(z)-f(w)|^2 \leq (1+\epsilon)|z-w|^2 \qquad (5\text{-}7)$$

The above calculation also can be achieved from **Formula 5-8** as follows [122]:

$$\frac{|f(z)-f(w)|^2}{(1+\epsilon)} \leq |z-w|^2 \leq \frac{|f(z)-f(w)|^2}{(1-\epsilon)} \qquad (5\text{-}8)$$

As demonstrated in **Formula 5-8**, the distance between the set of points in the lower-dimension space is nearly close to the distance in high-dimensional space. The Lemma theory asserts that it

is viable to project a set of points from a high-dimensional space into a lower-dimensional space, as the distances between the points are roughly preserved.

As a result, the above analysis proposes that if the initial set of features are projected into space with a lower-dimensional subspace using the random projection method, the distances between points are maintained under better contrast. Thus, it may enhance the classification accuracy between the features of two classes representing cases either with or without PM under low risk of overfitting ML models.

In this study, I also explore whether using RPA can yield a better result in comparison to several commonly used feature dimensionality reduction methods applied in the medical imaging informatics field, such as principal component analysis (PCA) [123], maximum relevance, and minimum redundancy (MRMR) [124], recursive feature elimination (RFE) [125], and least absolute shrinkage and selection operator (LASSO) [126]. All features extracted in the above section are fed into different methods of RPA, PCA, LASSO, MRMR, and RFE. Each method produces 20 optimal features out of the initially large set of 315 features.

5.2.5 Developing Machine Learning Model

To classify the study cases with or without PM, I build a multi-feature fusion-based machine learning model. Nevertheless, due to the unbalance dataset, which contains 121 PM cases and 38 non-PM cases, I employ a synthetic minority oversampling technique (SMOTE) algorithm [127] to rebalance the original image dataset. The advantages of using SMOTE to develop machine learning models in medical images have been well explored and verified in many previous studies (including those conducted by researchers in our lab) [90, 128, 129]. In this study, I apply the

SMOTE method to create 83 synthetic non-PM cases. Hence, the dataset is enlarged to 242 cases, including 121 PM cases and 121 non-PM cases.

After dealing with the imbalance dataset, I choose and apply the Gradient Boosting Machine (GBM) to train an optimal machine learning model to predict the risk of advanced gastric cancer patients suffering from having PM. The GBM model is a popular machine learning algorithm that has proven efficient in classifying complex datasets and often first in class with predictive accuracy [130]. Regarding hyperparameter tuning, the GBM model is implemented to achieve a low computational cost and high robustness in detection results as well. Moreover, to reduce the case partition and feature selection (or generation) bias, I apply a leave-one-case-out (LOCO) based cross-validation method to train and test the GBM model. In each LOCO cycle, PRA and SMOTE are embedded in the training process. Afterward, one case not involved in the training cycle is tested by the GBM model trained using all other cases in the dataset. The model generates a prediction score for each testing case ranging from 0 to 1. A higher score specifies a higher risk of PM. The prediction performance is assessed using a receiver operating characteristic (ROC) method after discarding all SMOTE generated non-PM training samples. The areas under ROC curves (AUC) and overall prediction accuracy after applying an operating threshold ($T = 0.5$) on the GBM model produced prediction scores are used as two performance evaluation indices. Furthermore, Cohen's Kappa coefficient value is also computed for evaluating the performance of the CAD scheme. High Cohen's Kappa coefficient value (ranging from zero to one) explains high robustness and less randomness in the predicted results [41, 131].

In summary, **Figure5-5** displays a complete flow chart using our CAD scheme to process images, compute optimal features, and train the GBM model in which the RPA and SMOTE are embedded inside the LOCO process. In this study, the segmentation and feature extraction steps

were performed using the MATLAB R2019a package, and the feature reduction and classifications were made using Python 3.7.

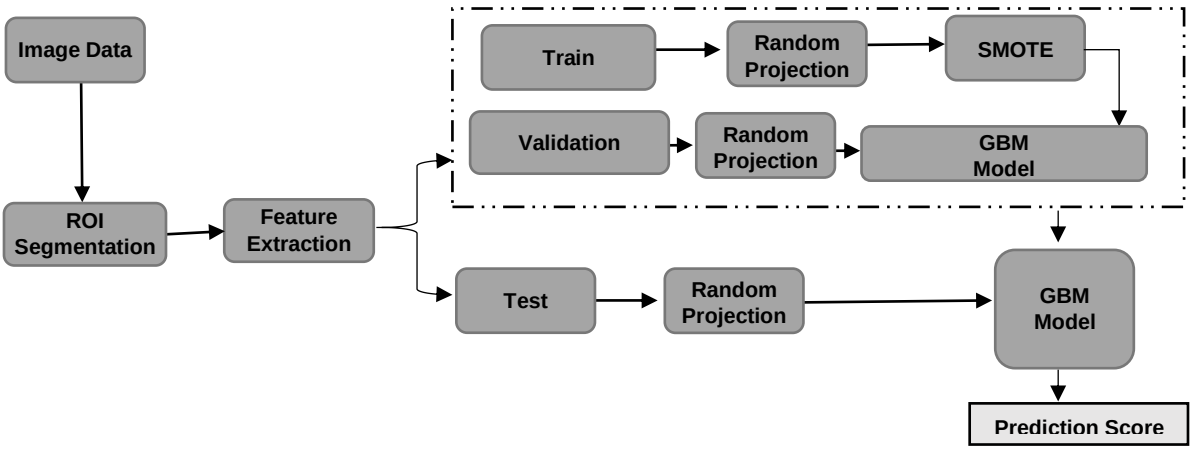


Figure 5-5. The flowchart of the CAD scheme proposed in this study.

5.3 Results

Figure 5-6 presents five ROC curves created by the GBM models embedded with five different feature reduction methods (LASSO, PCA, RFE, RPA, MRMR). **Table 5-2** illustrates the performance comparison between using RPA and the other four feature selection methods. The AUC value and the overall prediction accuracy of the GBM model trained using RPA with 3D image features as input are 0.69±0.019 and 71.2%, respectively. Besides, the precision, sensitivity, and specificity of the proposed method are 65.78%, 43.101%, and 87.12%, respectively. The results show that using RPA leads to produce an optimal image feature vector that can build a GBM model with substantially higher prediction accuracy ($p < 0.05$) than using the GBM models optimized using other four feature optimization methods.

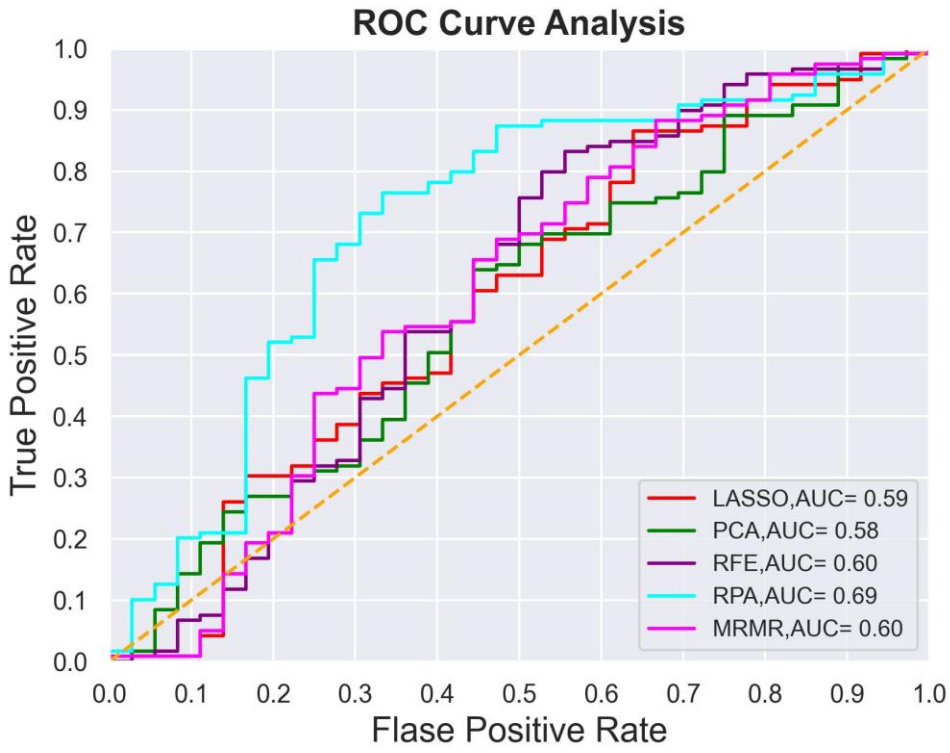


Figure 5-6. Comparison of five different ROC plots generated using GBM models optimized using five feature reduction and selection methods.

Table 5-2. The performance evaluation of five GBM models optimized using five different feature reduction and selection methods.

	Precision	Sensitivity	Specificity	Accuracy	AUC
LASSO	38.9%	31.1%	80.0%	65.8%	0.59±0.013
PCA	38.5%	64.1%	65.5%	65.2%	0.58±0.021
RFE	56.5%	62.5%	51.2%	56.9%	0.60±0.020
MRMR	50.0%	32.7%	82.0%	64.5%	0.60±0.017
RPA	65.8%	43.1%	87.1%	71.2%	0.69±0.019

Figure 5-7 demonstrates two ROC curves, and **Table 5-3** presents the prediction performance values to compare two GBM models trained using 2D features computed from the largest tumor region segmented from one CT image slice and the 3D features computed from the segmented tumor volumes. In these two GBM models, the RPA method is used to choose and generate optimal feature vectors. The results show that using 3D image features yields considerably higher performance than using 2D features ($p < 0.05$) in predicting the risk of gastric cancer cases with PM.

Table 5-3. The comparison of two GBM model performances between using 2D and 3D image features produced using the RPA method.

	AUC	**Accuracy**
2D features	0.66±0.017	68.4%
3D features	0.69±0.019	71.2%

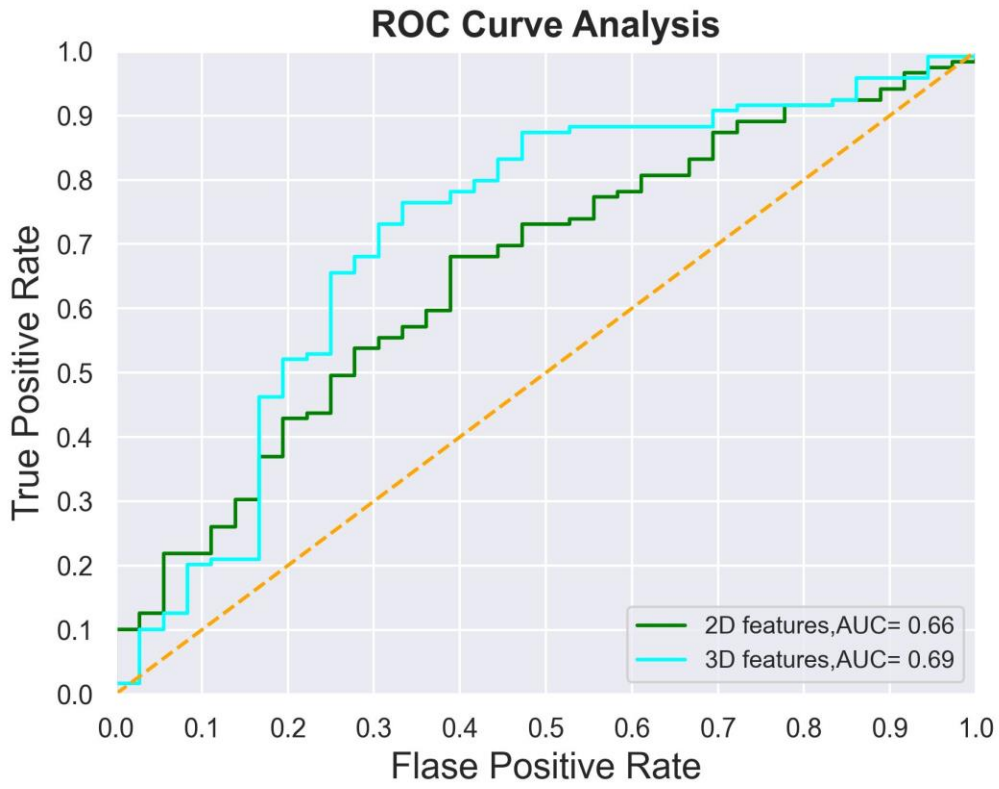


Figure 5-7. The comparison of two ROC plots produced by two GBM models optimized using 2D and 3D features generated using the RPA method, respectively.

Additionally, I also build and compare several other types of ML models, consisting of logistic regression, random forest, support vector machine (SVM), and decision tree. All models are trained and tested by the same LOCO cross-validation method embedded with RPA and SMOTE schemes. **Table 5-4** and **Figure 5-8** show the results to compare the prediction performance of five different ML models, which indicates that GBM yields the highest accuracy than the other four ML models. However, AUC values between GBM, SVM, and logistic regression-based ML models are not statistically significantly different ($p > 0.05$).

Table 5-4. The comparison of prediction performance of five different ML models.

	AUC value	Accuracy
SVM	0.66	64.55%
Logistic Regression	0.68	61.93%
Random Forest	0.63	69.03%
Decision Tree	0.56	65.16%
GBM	0.69	71.15%

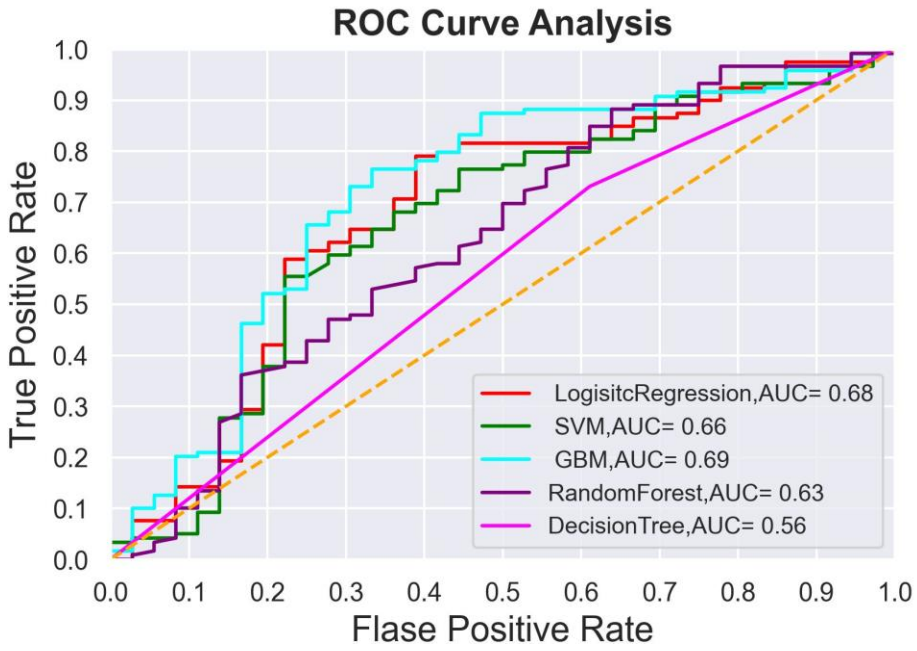


Figure 5-8. The comparison of ROC plots of five different ML models.

5.4 Discussion

CT is the most popular imaging modality to detect and diagnose different types of abnormality, such as gastric cancer, and it may also offer a non-invasive alternative method to predict the risk of PM in advanced gastric cancer patients. Regardless of the potential advantages of using CT to detect or predict the risk of PM, the effectiveness of radiologists in reading and interpreting CT images for PM detection is inadequate [132]. While studies have proposed that developing and applying CAD schemes combined with the radiomics concept and ML model is useful and may offer radiologists a second opinion to more accurately detect and diagnose different abnormalities [5], developing ML models using a large pool of radiomics features and small training dataset remains an arduous task. In this study, I examine a new approach to develop a new CAD scheme or ML model with several exceptional characteristics and novel ideas in feature extraction and ML model optimization to enhance accuracy in detecting advanced gastric patients with PM.

First, in a previous study, the authors carried out the manual segmentation of gastric cancer tumor regions from the single CT image slices [133]. Nevertheless, manual segmentation of tumor regions is often inconsistent with large inter-observer variability due to the fuzzy boundary of the tumor regions, which makes the computed image features also unreliable or not reproducible. Therefore, the prediction accuracy may be affected or not robust. To address this issue, in our study, I developed a new interactive CAD scheme with a graphical user interface (GUI) to initiate the segmentation of tumor regions from CT images. A user only requires placing an initial seed around the center of the tumor region with the largest area size in one CT slice. The CAD scheme then segments tumor regions on all involved CT image slices automatically. The segmentation results can also be visualized by the human eyes on the GUI window. Although I have designed and installed a correction function in the GUI and the operator can activate this function to order

CAD scheme correcting the segmentation errors (if any), the results in this study demonstrate that the CAD scheme can achieve satisfactory results on automatic segmentation of all 3,305 tumor regions from all 159 cases in our dataset.

Second, although the previous study [134] has suggested developing a radiomics based ML model to detect and diagnose gastric cancer using CT images, in that study, the Authors employ image features computed just from one manual selected CT image slice, which may not accurately represent image features of the entire tumor. To solve this issue, I conduct the first study that develops and tests a new ML model using 3D extracted image features. Our study results support our hypothesis that using 2D image features extracted from only one CT slice is not adequate enough to represent the heterogonous characteristics of the tumors while applying 3D image features can yield significantly higher prediction performance. In particular, in this study, I have performed 3D tumor segmentation and extracted 3D image features to detect or predict advanced gastric patients' risk of having PM. As revealed in **Table 5-3**, the prediction performance of the GBM model trained using 3D features yields AUC=0.69±0.019 and an accuracy of 71.2%, which are notably higher than the GBM model trained using 2D features with AUC=0.66±0.017 and the accuracy of 68.4% ($p < 0.05$), respectively.

Third, in developing CAD schemes to train ML models, identifying a small and effective set of image features plays an important role [135, 136]; thus, in previous studies, different feature dimensionality reduction methods have been investigated [137, 138]. Although these studies made noticeable progress in optimizing the feature vectors, there is a substantial challenge of achieving small feature vectors representing the complex and non-linear image feature space. In this study, I explore the viability of applying the RPA to the medical imaging informatics field in optimizing the CAD scheme or ML model. Our study results reveal that RPA is a promising technique to

reduce the dimensionality of a set of points lying in Euclidian space for very heterogeneous feature data, which commonly occurs in medical images and has advantages to attain high robustness in classification and low risk of overfitting. **Figure 5-6** shows that the prediction performance of the GBM model embedded with RPA yields substantially higher performance than other GBM models embedded with the other four popular feature reduction methods (PCA, LASSO, MRMR, and RFE). As presented in **Table 5-2**, the AUC value after applying the RPA reached the highest prediction accuracy of 71.2% than the other four feature reduction methods. Moreover, the computed Cohen's Kappa coefficient value is 0.68, which specifies the reliability or robustness of the GBM model optimized using the RPA method.

Fourth, since many ML models have been developed and used in medical imaging informatics or CAD fields, choosing which ML model can also be a challenging issue. In this study, I also evaluate the prediction performance of five popular ML models. The results demonstrate that many different ML models can yield very comparable performance, as shown in Table 4 and Figure 8. However, considering the data presented in **Table 5-2**, I can find that selecting or generating optimal features plays a more essential role or contribution than choosing a different ML model. Thus, combining the above new observations of this study, I validate that due to the very complicated distribution of radiomics features extracted from medical images, RPA is a promising and more powerful technique applicable to generate optimal feature vectors for better training ML models used in CAD schemes of medical images.

This is a valid proof-of-concept study that indicates a new and promising approach to identify and generate optimal feature vectors for training ML models implemented in CAD schemes of medical images. Since optimizing the feature vector is one of the essential steps of building an optimal ML model using the radiomics concept, the proposed method in this study is not only

limited to the detection of advanced gastric patients with PM, and it can also be beneficial for other medical imaging studies of developing ML models to detect different types of cancers or abnormalities in the future studies.

Chapter 6: Summary and Future Work

6.1 Summary

Computer-aided detection schemes have been developed and are currently widely employed as a second reader in many medical facilities.

In the previous chapters of this dissertation, I presented different studies that I have done during my Ph.D. In chapter 1, I have discussed the CAD concept, application, and also challenges in developing current CAD schemes. In chapter 2, I expressed the overall objective of this dissertation as well as each of my studies' hypotheses and motivations.

As discussed in chapter 3, my first study's purpose was to use the false-positive generated by the CAD schemes to predict the short-term breast cancer risk [46]. In that study, I hypothesized that the false positives generated by the CAD schemes are not harmful and may be beneficial to extract some useful information to predict the cancer risk. To test that hypothesis, I have used a relatively large dataset and used a previously developed CAD scheme in our lab to extract new quantitative features from the CAD-generated false positive to predict the short-term breast cancer risk. The study showed the promising results of using the new extracted features in breast cancer risk prediction.

Following that study, In chapter 4, I have presented my second study for predicting the treatments' efficacy in reducing cancer tumor size [113]. In that study, I hypothesis that the image features extracted from the ultrasound images could be valuable in evaluating the treatments' performance in reducing the tumor size. For this study, an image dataset involving ultrasound images of 23 athymic nude mice bearing C26 mouse adenocarcinoma cell tumors was assembled. These mice were divided into seven groups and treated using seven thermal therapy methods,

respectively. Additionally, I have developed a CAD scheme to conduct tumor segmentation and feature analysis. The study demonstrated the feasibility of extracting new quantitative image features from ultrasound images to assist the early assessment of tumor response to therapies.

In Chapter 5, I have reported my last study [139] in investigating a new method to optimize the ML model for predicting PM in gastric cancer patients. This study's hypothesis was that by applying a new feature reduction method named RPA, the ML model's accuracy in predicting PM could be enhanced. We assemble a retrospective dataset involving abdominal computed tomography (CT) images acquired from advanced gastric cancer patients. A computer-aided detection (CAD) scheme is then developed and applied to segment the gastric tumor area and computes image features. The study showed that CT images of the gastric tumors contain discriminatory information to predict PM's risk in gastric cancer patients, and RPA is a promising method to generate optimal feature vectors to improve ML models' performance.

6.1.1 The Contributions of My Ph.D. Research Work

1- I introduced new quantitative image features and used the false-positive results generated by the CAD schemes. Due to this study's novelty and unique characteristics, it was selected as a featured article in the journal of Physics and Medicine in Biology in 2018.

2- I extracted quantitative features from ultrasound images to predict treatment’s efficacy, which has not been investigated before my study.

3- I applied a new feature reduction method to select the optimal features from the large initial pool of features for the first time in the medical imaging field.

4- I proposed a new method to segment the cancer tumor volume automatically.

5- I developed several interactive CAD schemes with a graphic user interface to increase transparency, assist users in better understanding CAD schemes' results, and enhance users' confidence (clinicians) to consider and accept CAD-generated results.

Through my studies, we have published several journals and conference papers, and peer-reviewed abstracts, as shown in the following section in which I was either first or co-author.

In Summary, in addition to the publications, one of my research work's significant contributions is that I have collaborated with clinicians in OUHSC and developed CAD schemes with GUI by which they developed promising cancer studies, and their results have been reported in independent studies. Hence, my research studies demonstrated the feasibility of developing novel quantitative imaging analysis schemes based on machine learning for cancer research.

6.1.2 Journal Papers

1. Buechel, M.E., Enserro, D., Burger, R.A., Brady, M.F., Wade, K., Secord, A.A., Nixon, A.B., **Mirniaharikandehei, S.**, Liu, H., Zheng, B. and O'Malley, D.M., (2021). Correlation of imaging and plasma based biomarkers to predict response to bevacizumab in epithelial ovarian cancer (EOC). Gynecologic oncology.
2. **Mirniaharikandehei, S**., Heidari, M., Danala, G., Lakshmivarahan, S., and Zheng, B. (2021). Applying a random projection algorithm to optimize machine learning model for predicting peritoneal metastasis in gastric cancer patients using CT images. Computer Methods and Programs in Biomedicine, 200,105937.
3. Heidari, M., Lakshmivarahan, S., **Mirniaharikandehei, S.**, Danala, G., Maryada, k., Liu. H., and Zheng, B., (2021). Applying a random projection algorithm to optimize machine learning model for breast lesion classification. IEEE Transactions on Biomedical Engineering, 10.1109.
4. Heidari, M., **Mirniaharikandehei, S.**, Khuzani, A. Z., Danala, G., Qiu, Y., and Zheng, B., (2020). Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. International Journal of Medical Informatics, 144, 104284.

5. Zheng, B., Qiu, Y., Aghaei, F., **Mirniaharikandehei, S.**, Heidari, M. and Danala, G., (2019). Developing global image feature analysis models to predict cancer risk and prognosis. Visual Computing for Industry, Biomedicine, and Art, 2(1), 1-14.

6. Heidari, M., **Mirniaharikandehei, S.**, Liu, W., Hollingsworth, A. B., Liu, H., and Zheng, B., (2019). Development and Assessment of a New Global Mammographic Image Feature Analysis Scheme to Predict Likelihood of Malignant Cases. IEEE transactions on medical imaging, 39(4), 101109.

7. **Mirniaharikandehei, S.**, VanOsdol, J., Heidari, M., Danala, G., Sethuraman, S.N., Ranjan, A. and Zheng, B., (2019). Developing a quantitative ultrasound image feature analysis scheme to assess tumor treatment efficacy using a mouse model. Scientific reports, 9(1), 1-10.

8. **Mirniaharikandehei, S.,** Hollingsworth, A. B., Patel, B., Heidari, M., Liu, H., and Zheng, B., (2018). Applying a new computer-aided detection scheme generated imaging marker to predict short-term breast cancer risk. Physics in Medicine & Biology, 63(10), 105005**.**

9. Heidari, M., Khuzani, A. Z., Hollingsworth, A. B., Danala, G., **Mirniaharikandehei, S.**, Qiu, Y., and Zheng, B. (2018). Prediction of breast cancer risk using a machine learning approach embedded with a locality preserving projection algorithm. Physics in Medicine and Biology, 63(3), 035020-035020.

6.1.3 Conference Proceeding Papers

1. **Mirniaharikandehei, S.**, Heidari, M., Danala, G., Lakshmivarahan, S., and Zheng, B. (2021, February). A novel feature reduction method to improve performance of machine learning model. In Medical Imaging 2021: Computer-Aided Diagnosis (Vol. 11597, p. 1159726). International Society for Optics and Photonics.

2. Heidari, M., **Mirniaharikandehei, S.**, Khuzani, A. Z. , Danala, G., Pham, H., Lakshmivarahan, S., and Zheng, B. (2021, February). An optimal machine learning model for breast lesion classification based on random projection algorithm for feature optimization. In Medical Imaging 2021: Imaging Informatics for Healthcare, Research, and Applications (Vol. 11601, p. 116010L). International Society for Optics and Photonics.

3. Heidari, M., **Mirniaharikandehei, S.**, Abolfazl Zargari Khuzani, Danala, G., Qiu, Y., and Zheng, B. (2021, February). Detecting COVID-19 infected pneumonia from x-ray images using a deep learning model with image preprocessing algorithm. In Medical Imaging 2021: Computer-Aided Diagnosis, Research, and Applications (Vol. 11597, p. 115970V). International Society for Optics and Photonics.

4. Heidari, M., **Mirniaharikandehei, S.**, Danala, G., Qiu, Y. and Zheng, B., (2020). A new case-based CAD scheme using a hierarchical SSIM feature extraction method to classify between malignant and benign cases. In SPIE (Vol. 11318, p. 1131816).

5. Heidari, M., Hollingsworth, A., **Mirniaharikandehei, S**., Danala, G., Qiu, Y., Liu, H., and Zheng, B., (2019). Assessment of short-term breast cancer risk using a frequency domain correlation based imaging marker. In Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications (Vol. 10954, p. 109541F). International Society for Optics and Photonics.

6. Zarafshani, A., Wang, Y., **Mirniaharikandehei, S**.N., Heidari, M., Aghaei, F., Wang, S., Xiang, L. and Zheng, B., (2019). Design, fabrication and evaluation of non-imaging, label-free pre-screening tool using quantified bio-electrical tissue profile. In SPIE (Vol. 10953, p. 1095304).

7. Aghaei, F., Hollingsworth, A. B., **Mirniaharikandehei, S**., Wang, Y., Liu, H., and Zheng, B., (2019). Developing a new quantitative imaging marker to predict pathological complete response to neoadjuvant chemotherapy. In Medical Imaging 2019: Computer-Aided Diagnosis (Vol. 10950, p. 109502O). International Society for Optics and Photonics.

8. **Mirniaharikandehei, S**., Heidari, M., Danala, G., Qian, W., Qiu, Y., and Zheng, B., (2019). Association of computer-aided detection results and breast cancer risk. In Medical Imaging 2019: Computer-Aided Diagnosis (Vol. 10950, p. 109500I). International Society for Optics and Photonics.

9. **Mirniaharikandehei, S.**, VanOsdol, J., Heidari, M., Danala, G., Ranjan, A. and Zheng, B., (2019). Developing a quantitative ultrasound image feature analysis scheme to assess tumor treatment efficacy using a mouse model. In Medical Imaging 2019: Ultrasonic Imaging and Tomography (Vol. 10955, p. 109551A). International Society for Optics and Photonics.

10. Heidari, M., **Mirniaharikandehei, S.**, Khuzani, A. Z., Qian, W., Qiu, Y., and Zheng, B., (2019). Assessment of a quantitative mammographic imaging marker for breast cancer risk prediction. In Medical Imaging 2019: Image Perception, Observer Performance, and Technology Assessment (Vol. 10952, p. 109520X). International Society for Optics and Photonics.

11. Aghaei, F., **Mirniaharikandehei, S.**, Hollingsworth, A. B., Stoug, R. G., Pearce, M., Liu, H., and Zheng, B, (2018). Association between background parenchymal enhancement of breast MRI and BIRADS rating change in the subsequent screening. In Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications (Vol. 10579, p. 105790R). International Society for Optics and Photonics.

12. Wang, Y., Heidari, M., **Mirniaharikandehei, S.**, Gong, J., Qian, W., Qiu, Y., and Zheng, B., (2018). A hybrid deep learning approach to predict malignancy of breast lesions using mammograms. In Medical Imaging 2018: Imaging Informatics for

Healthcare, Research, and Applications (Vol. 10579, p. 105790V). International Society for Optics and Photonics.

13. Zarafshani, A., Dhurjaty, S., **Mirniaharikandehei, S. N.**, Aghaei, F., Xiang, L., and Zheng, B., (2018). Developing a unique portable device to non-invasively detect bio-electrochemical characteristics of human tissues. In Medical Imaging 2018: Physics of Medical Imaging (Vol. 10573, p. 105735L). International Society for Optics and Photonics.

14. Heidari, M., Khuzani, A. Z., Danala, G., **Mirniaharikandehei, S.**, Qian, W., and Zheng, B. ,(2018). Applying a machine learning model using a locally preserving projection-based feature regeneration algorithm to predict breast cancer risk. In Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications (Vol. 10579, p. 105790T). International Society for Optics and Photonics.

15. **Mirniaharikandehei, S.**, Zarafshani, A., Heidari, M., Wang, Y., Aghaei, F., and Zheng, B., (2018). Applying a CAD-generated imaging marker to assess short-term breast cancer risk. In Medical Imaging 2018: Computer-Aided Diagnosis (Vol. 10575, p. 105753F). International Society for Optics and Photonics.

16. **Mirniaharikandehei, S.**, Patil, O., Aghaei, F., Wang, Y., and Zheng, B. ,(2017). Exploring a new quantitative image marker to assess benefit of chemotherapy to ovarian cancer patients. In Medical Imaging 2017: Imaging Informatics for Healthcare, Research, and Applications (Vol. 10138, p. 101380I). International Society for Optics and Photonics.

17. Aghaei, F., **Mirniaharikandehei, S**., Hollingsworth, A. B., Wang, Y., Qiu, Y., Liu, H., and Zheng, B. (2017). Exploring a New Bilateral Focal Density Asymmetry Based Image Marker to Predict Breast Cancer Risk. In SPIE Medical Imaging (pp. 101361P-101361P). International Society for Optics and Photonics.

6.1.4 Published Abstract

1. Buechel, M., Enserro, D., Burger, R.A., Brady, M.F., Wade, K., Secord, A.A., Nixon, A.B., **Mirniaharikandehei, S.**, Liu, H., Zheng, B. and Gray, H., (2018). Correlation of imaging and plasma-based biomarkers to predict response to bevacizumab in epithelial ovarian cancer (EOC): A GOG 218 ancillary data analysis.

2. Buechel, M.E., Brady, M.F., Wang, Y., **Mirniaharikandehei, S.**, Liu, H., Zheng, B., Salani, R., Tewari, K.S., Gray, H.J., Bakkum-Gamez, J.N. and Burger, R.A., (2017). Measurements of sarcopenia as predictive biomarkers for response to anti-angiogenic treatment in epithelial.

6.2 Future Studies

Despite the encouraging results of recent studies in imaging analysis schemes, it still suffers from some limitations that needed to be explored in future studies. For instance, most developed schemes are suffering from proper and highly accurate automatic segmentation of the region of interest, which leads to inaccurate feature selection and finally affects the classifications' accuracy. Apart from the segmentation step, as previously mentioned, other imaging analysis schemes such as feature extraction, feature selection, and classification are required to be improved. Hence, these areas have been attracting many research interests and need further investigations in future studies. Furthermore, in the following paragraphs, I want to briefly mention some limitations in each of my studies which can be addressed and evaluated in future studies.

Although my first study had several unique characteristics, it also had some limitations. First, while I used a relatively large dataset involving 1,044 mammography screening cases, it has a very high cancer prevalence ratio (402 vs. 642), which does not represent the cancer prevalence ratio in the real mammography screening environment. Instead, this is just a laboratory-based retrospective study. This approach's potential clinical utility and/or new imaging marker need to be validated in future prospective studies. Second, this study only focuses on predicting the risk of developing mammography-detectable cancer in the subsequent screening. Although detection or distinction between invasive (or aggressive) and non-invasive cancer is an important clinical issue and research topic in reducing over-diagnosis and over-treatment, whether we can develop a new clinical marker or model based on CAD of mammograms to help solve this issue has not been investigated in this study. Third, I only extracted four simple features from the CAD-generated detection results and selected a simple multinomial logistic regression model. To further advance

risk prediction performance, more studies are needed to continue exploring, identifying, or developing more effective imaging markers from the CAD-generated detection results and their impact on future breast cancer risk.

As for my second study, despite the promising data analysis results and observations, this was a preliminary study with limitations. For example, Althoughin this study, I only used a small dataset including 23 athymic nude mice, which were divided into seven groups treated with seven different thermal therapies. The ultrasound images were acquired by one group of researchers in one research laboratory. Thus, in order to validate the study results and enhance the feasibility of developing robust image feature based markers to predict cancer treatment efficacy, more studies are needed by using larger and diverse datasets from both prior clinical (using mouse models) and clinical (using real patient images) researches in the future.

Regardless of my last study's encouraging results, I also notice some limitations in the presented study. First, the dataset used in my study is moderately small; therefore, to validate this study's results, larger datasets are essential before being tested in future prospective clinical studies. Second, despite the fact that in this study, I have used synthetic data to balance the dataset and reduce the impact of an imbalanced dataset, applying the SMOTE technique is just sufficient for the low dimensional data, and it may not be appropriate or optimal for high dimensional data [140]. Third, in the large initial set of features, I only extracted a limited number of 315 statistics and textural features, which are much less than the number of features extracted and computed based on recently developed radiomics concepts and technology in other studies [141]. Hence, more texture features can be investigated in future studies to enhance the diversity of the initial feature pool, which may also increase the chance of selecting or generating more optimal features

to substantially improve the ML model's accuracy to predict the risk of PM. To address the above limitations, more studies and progress are needed in this field.

Appendix I

7.1 The Gray Level Co-Occurrence Matrixes (GLCM) Features

Texture features were calculated using the GLCM. The results are nine texture parameters. Suppose GLCM is defined as Cij. The angular second moment, inertia, inverse difference moment, entropy, correlation, sum average, difference average, sum entropy, Contrast, the sum of squares variance, sum variance, difference variance, and difference entropy are computed as follows:

1- The angular second moment

$$T_1 = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} C_{ij}^2$$

2- Inertia

$$T_2 = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} (i-j)^2 C_{ij}$$

3- Inverse difference moment

$$T_3 = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \frac{1}{1+(i-j)^2} C_{ij}$$

4- Entropy

$$T_4 = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} C_{ij} \log C_{ij}$$

5- Correlation

$$T_5 = \frac{\sum_{i=0}^{k-1}\sum_{j=0}^{k-1}(ij)C_{ij} - \mu_x\mu_y}{\sigma_x\sigma_y}$$

6- Sum average

$$T_6 = \sum_{k=0}^{2k-2} kC_{x+y}(k)$$

7- Difference average

$$T_7 = \sum_{k=0}^{k-1} kC_{x-y}(k)$$

8- Sum entropy

$$T_8 = -\sum_{k=0}^{2k-2} C_{x+y}(k) \log\{C_{x+y}(k)\}$$

9- Difference entropy

$$T_9 = -\sum_{k=0}^{k-1} C_{x-y}(k) \log\{C_{x-y}(k)\}$$

10- Contrast

$$T_{10} = \sum_{i=1}^{N-1} n^2 \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} \{p(i,j)\} \right\}$$

11- Sum of squares variance

$$T_{11} = \sum_{i} \sum_{j} (i-\mu)^2 C_{ij}$$

12- Sum variance

$$T_{12} = \sum_{i=2}^{2N} (i-t_s)^2 C_{(x+y)}(i,j)$$

13- Difference variance

$$T_{13} = variance\ of\ C_{(x-y)}$$

7.2 Gray Level Run Length Matrix-based features (GLRLM)

Gray-level run-length matrix-based features, as defined by [142]. Assume (i,j)th entry in the given run-length matrix is P(i,j), the number of different run-length is N_r, the number of discrete intensity value in the image is N_g, the number of voxels in the image is N_p, and the entry (i,j) of the normalized GLRLM is described as:

$$p(i,j) = \frac{p(i,j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)}$$

1- Short Run Emphasis (SRE)

$$SRE = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left[\frac{p(i,j)}{j^2} \right]$$

2- Long Run Emphasis (LRE)

$$LRE = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 p(i,j)$$

3- Gray level Non-Uniformity (GLN)

$$GLN = \sum_{i=1}^{N_g} \left[\sum_{j=1}^{N_r} p(i,j) \right]^2$$

4- Run Length Non-Uniformity (RLN)

$$RLN = \sum_{i=1}^{N_r} \left[\sum_{j=1}^{N_g} p(i,j) \right]^2$$

5- Run Percentage (RP)

$$RP = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{p(i,j)}{N_p}$$

6- Low Gray Level Run Emphasis (LGRE)

$$LGRE = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left[\frac{p(i,j)}{i^2} \right]$$

7- High Gray Level Run Emphasis (HGRE)

$$HGRE = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} i^2 p(i,j)$$

8- Short Run Low Gray Emphasis (SRLGE)

$$SRLGE = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left[\frac{p(i,j)}{i^2 j^2} \right]$$

9- Short Run High Gray Emphasis (SRHGE)

$$SRHGE = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left[\frac{p(i,j) i^2}{j^2} \right]$$

10- Long Run Low Gray Emphasis (LRLGE)

$$LRLGE = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \left[\frac{p(i,j) j^2}{i^2} \right]$$

11- Long Run High Gray Emphasis (LRHGE)

$$LRHGE = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} i^2 j^2 p(i,j)$$

References

1. Lodwick, G.S., et al., *Computer diagnosis of primary bone tumors: a preliminary report.* Radiology, 1963. **80**(2): p. 273-275.
2. Meyers, P.H., et al., *Automated computer analysis of radiographic images.* Radiology, 1964. **83**(6): p. 1029-1034.
3. Winsberg, F., et al., *Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis.* Radiology, 1967. **89**(2): p. 211-215.
4. Doi, K., *Computer-aided diagnosis in medical imaging: historical review, current status and future potential.* Computerized medical imaging and graphics, 2007. **31**(4-5): p. 198-211.
5. Gonçalves, V.M., M.E. Delamaro, and F.d.L.d.S. Nunes, *A systematic review on the evaluation and characteristics of computer-aided diagnosis systems.* Revista Brasileira de Engenharia Biomédica, 2014. **30**(4): p. 355-383.
6. Heidari, M., et al. *Improving performance of breast cancer risk prediction using a new CAD-based region segmentation scheme*. in *Medical Imaging 2018: Computer-Aided Diagnosis*. 2018. International Society for Optics and Photonics.
7. Kumar, V., et al., *Radiomics: the process and the challenges.* Magnetic resonance imaging, 2012. **30**(9): p. 1234-1248.
8. Aerts, H., *L, Velazquez ER, Leijenaar RT H, Parmar C, Grossmann P, Carvalho S et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.* Nature Communications, 2014. **5**: p. 4006.

9. Tan, M., J. Pu, and B. Zheng, *A new and fast image feature selection method for developing an optimal mammographic mass detection scheme.* Medical physics, 2014. **41**(8Part1): p. 081906.

10. Erickson, B.J., et al., *Machine learning for medical imaging.* Radiographics, 2017. **37**(2): p. 505-515.

11. Camlica, Z., H.R. Tizhoosh, and F. Khalvati. *Medical image classification via svm using lbp features from saliency-based folded data*. in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*. 2015. IEEE.

12. Konukoglu, E. and B. Glocker, *Random forests in medical image computing*, in *Handbook of Medical Image Computing and Computer Assisted Intervention*. 2020, Elsevier. p. 457-480.

13. Ramteke, R. and K.Y. Monali, *Automatic medical image classification and abnormality detection using k-nearest neighbour.* International Journal of Advanced Computer Research, 2012. **2**(4): p. 190.

14. Gonçalves, V.M., M.E. Delamaro, and F.d.L.d.S. Nunes, *A systematic review on the evaluation and characteristics of computer-aided diagnosis systems.* Revista Brasileira De Engenharia Biomedica, 2014. **30**(4): p. 355-383.

15. Freer, T.W. and M.J. Ulissey, *Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center.* Radiology, 2001. **220**(3): p. 781-786.

16. Gur, D., et al., *Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system.* Journal of the National Cancer Institute, 2004. **96**(3): p. 185-190.

17. Birdwell, R.L., P. Bandodkar, and D.M. Ikeda, *Computer-aided detection with screening mammography in a university hospital setting.* Radiology, 2005. **236**(2): p. 451-457.

18. Morton, M.J., et al., *Screening mammograms: interpretation with computer-aided detection—prospective evaluation.* Radiology, 2006. **239**(2): p. 375-383.

19. Cupples, T.E., J.E. Cunningham, and J.C. Reynolds, *Impact of computer-aided detection in a regional screening mammography program.* American Journal of Roentgenology, 2005. **185**(4): p. 944-950.

20. Doi, K. *Overview on research and development of computer-aided diagnostic schemes*. in *Seminars in Ultrasound, CT and MRI*. 2004. Elsevier.

21. Sittek, H., et al., *Computer-assisted analysis of mammograms in routine clinical diagnosis.* Der Radiologe, 1998. **38**(10): p. 848-852.

22. Funovics, M., et al., *Computer-assisted diagnosis in mammography: the R2 ImageChecker system in detection of spiculated lesions.* Wiener medizinische Wochenschrift (1946), 1998. **148**(14): p. 321-324.

23. Fenton, J.J., et al., *Influence of computer-aided detection on performance of screening mammography.* New England Journal of Medicine, 2007. **356**(14): p. 1399-1409.

24. Nishikawa, R., et al. *Computer-aided diagnosis in screening mammography: Detection of missed cancers*. in *Radiology*. 1998. RADIOLOGICAL SOC NORTH AMER 20TH AND NORTHAMPTON STS, EASTON, PA 18042 USA.

25. Nawano, S., et al., *Computer-aided diagnosis in full digital mammography.* Investigative radiology, 1999. **34**(4): p. 310.

26. Nishikawa, R.M. and A. Bandos. *Predicting the benefit of using CADe in screening mammography*. in *International Workshop on Digital Mammography*. 2014. Springer.

27. Francis, S.V., M. Sasikala, and S. Saranya, *Detection of breast abnormality from thermograms using curvelet transform based feature extraction.* Journal of medical systems, 2014. **38**(4): p. 1-9.

28. Kriege, M., et al., *Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition.* New England Journal of Medicine, 2004. **351**(5): p. 427-437.

29. Nishikawa, R.M. and D. Gur, *CADe for early detection of breast cancer—Current status and why we need to continue to explore new approaches.* Academic radiology, 2014. **21**(10): p. 1320-1321.

30. Amir, E., et al., *Assessing women at high risk of breast cancer: a review of risk assessment models.* JNCI: Journal of the National Cancer Institute, 2010. **102**(10): p. 680-691.

31. Lian, J. and K. Li, *A review of breast density implications and breast cancer screening.* Clinical breast cancer, 2020.

32. McCann, J., D. Stockton, and S. Godward, *Impact of false-positive mammography on subsequent screening attendance and risk of cancer.* Breast Cancer Research, 2002. **4**(5): p. 1-9.

33. Castells, X., et al., *Breast cancer detection risk in screening mammography after a false-positive result.* Cancer epidemiology, 2013. **37**(1): p. 85-90.

34. Dalecki, D., K.P. Mercado, and D.C. Hocking, *Quantitative ultrasound for nondestructive characterization of engineered tissues and biomaterials.* Annals of biomedical engineering, 2016. **44**(3): p. 636-648.

35. Shu, Z., et al., *Prediction of efficacy of neoadjuvant chemoradiotherapy for rectal cancer: the value of texture analysis of magnetic resonance images.* Abdominal Radiology, 2019. **44**(11): p. 3775-3784.

36. Hou, Z., et al., *Radiomic analysis in contrast-enhanced CT: predict treatment response to chemoradiotherapy in esophageal carcinoma.* Oncotarget, 2017. **8**(61): p. 104444.

37. Bera, K., V. Velcheti, and A. Madabhushi, *Novel quantitative imaging for predicting response to therapy: techniques and clinical applications.* American Society of Clinical Oncology Educational Book, 2018. **38**: p. 1008-1018.

38. Aghaei, F., et al. *Developing a new quantitative imaging marker to predict pathological complete response to neoadjuvant chemotherapy*. in *Medical Imaging 2019: Computer-Aided Diagnosis*. 2019. International Society for Optics and Photonics.

39. Lambin, P., et al., *Radiomics: extracting more information from medical images using advanced feature analysis.* European journal of cancer, 2012. **48**(4): p. 441-446.

40. Zheng, B., et al., *Developing global image feature analysis models to predict cancer risk and prognosis.* Visual Computing for Industry, Biomedicine, and Art, 2019. **2**(1): p. 1-14.

41. Heidari, M., et al., *Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms.* International journal of medical informatics, 2020. **144**: p. 104284.

42. Heidari, M., et al. *A new case-based CAD scheme using a hierarchical SSIM feature extraction method to classify between malignant and benign cases*. in *Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications*. 2020. International Society for Optics and Photonics.

43. Mirniaharikandehei, S., et al. *Association of computer-aided detection results and breast cancer risk.* in *Medical Imaging 2019: Computer-Aided Diagnosis*. 2019. International Society for Optics and Photonics.

44. Gao, Y., et al., *New frontiers: an update on computer-aided diagnosis for breast imaging in the age of artificial intelligence.* American Journal of Roentgenology, 2019. **212**(2): p. 300-307.

45. Wang, Y., et al. *A hybrid deep learning approach to predict malignancy of breast lesions using mammograms*. in *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*. 2018. International Society for Optics and Photonics.

46. Mirniaharikandehei, S., et al., *Applying a new computer-aided detection scheme generated imaging marker to predict short-term breast cancer risk.* Physics in Medicine & Biology, 2018. **63**(10): p. 105005.

47. Mirniaharikandehei, S., et al. *Applying a CAD-generated imaging marker to assess short-term breast cancer risk.* in *Medical Imaging 2018: Computer-Aided Diagnosis*. 2018. International Society for Optics and Photonics.

48. Mirniaharikandehei, S., et al. *Exploring a new quantitative image marker to assess benefit of chemotherapy to ovarian cancer patients*. in *Medical Imaging 2017: Imaging Informatics for Healthcare, Research, and Applications*. 2017. International Society for Optics and Photonics.

49. Goel, N., G. Bebis, and A. Nefian. *Face recognition experiments with random projection*. in *Biometric Technology for Human Identification II*. 2005. International Society for Optics and Photonics.

50. Bingham, E. and H. Mannila. *Random projection in dimensionality reduction: applications to image and text data*. in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. 2001.

51. Xu, B. and G. Qiu. *Crowd density estimation based on rich features and random projection forest*. in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2016. IEEE.

52. Berlin, L. and F.M. Hall, *More mammography muddle: emotions, politics, science, costs, and polarization.* Radiology, 2010. **255**(2): p. 311-316.

53. Carney, P.A., et al., *Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography.* Annals of internal medicine, 2003. **138**(3): p. 168-175.

54. Hubbard, R.A., et al., *Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study.* Annals of internal medicine, 2011. **155**(8): p. 481-492.

55. Brodersen, J. and V.D. Siersma, *Long-term psychosocial consequences of false-positive screening mammography*. The Annals of Family Medicine, 2013. **11**(2): p. 106-115.

56. Brawley, O.W., *Risk-based mammography screening: an effort to maximize the benefits and minimize the harms.* Annals of internal medicine, 2012. **156**(9): p. 662-663.

57. Gail, M.H. and P.L. Mai, *Comparing breast cancer risk assessment models*. 2010, Oxford University Press.

58. Hollingsworth, A.B. and R.G. Stough, *An alternative approach to selecting patients for high-risk screening with breast MRI.* The breast journal, 2014. **20**(2): p. 192-197.

59. Wei, J., et al., *Association of computerized mammographic parenchymal pattern measure with breast cancer risk: a pilot case-control study.* Radiology, 2011. **260**(1): p. 42-49.

60. Zheng, B., et al., *Bilateral mammographic density asymmetry and breast cancer risk: a preliminary assessment.* European journal of radiology, 2012. **81**(11): p. 3222-3228.

61. Tan, M., et al., *Association between changes in mammographic image features and risk for near-term breast cancer development.* IEEE transactions on medical imaging, 2016. **35**(7): p. 1719-1728.

62. Astley, S. and F.J. Gilbert, *Computer-aided detection in mammography.* Clinical radiology, 2004. **59**(5): p. 390-399.

63. Birdwell, R.L., et al., *Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection.* Radiology, 2001. **219**(1): p. 192-202.

64. Zheng, B., et al., *Performance change of mammographic CAD schemes optimized with most-recent and prior image databases.* Academic radiology, 2003. **10**(3): p. 283-288.

65. Zheng, B., et al., *Computer-aided detection: the effect of training databases on detection of subtle breast masses.* Academic radiology, 2010. **17**(11): p. 1401-1408.

66. Zheng, B., et al., *Computer-aided detection of breast masses depicted on full-field digital mammograms: a performance assessment.* The British journal of radiology, 2012. **85**(1014): p. e153-e161.

67. Zheng, B., Y.-H. Chang, and D. Gur, *Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis.* Academic radiology, 1995. **2**(11): p. 959-966.

68. Gur, D., et al., *Computer-aided detection performance in mammographic examination of masses: assessment.* Radiology, 2004. **233**(2): p. 418-423.

69. Emaminejad, N., et al., *Fusion of quantitative image and genomic biomarkers to improve prognosis assessment of early stage lung cancer patients.* IEEE Transactions on Biomedical Engineering, 2015. **63**(5): p. 1034-1043.

70. Danala, G., et al., *Applying quantitative CT image feature analysis to predict response of ovarian cancer patients to chemotherapy.* Academic radiology, 2017. **24**(10): p. 1233-1239.

71. Heidari, M., et al., *Prediction of breast cancer risk using a machine learning approach embedded with a locality preserving projection algorithm.* Physics in Medicine & Biology, 2018. **63**(3): p. 035020.

72. Li, Q. and K. Doi, *Reduction of bias and variance for evaluation of computer-aided diagnostic schemes.* Medical physics, 2006. **33**(4): p. 868-875.

73. McCann, J., D. Stockton, and S. Godward, *Impact of false-positive mammography on subsequent screening attendance and risk of cancer.* Breast Cancer Research, 2002. **4**(5): p. R11.

74. Giess, C.S., E.P. Frost, and R.L. Birdwell, *Interpreting one-view mammographic findings: minimizing callbacks while maximizing cancer detection.* RadioGraphics, 2014. **34**(4): p. 928-940.

75. Wang, X., et al., *Computerized prediction of risk for developing breast cancer based on bilateral mammographic breast tissue asymmetry.* Medical engineering & physics, 2011. **33**(8): p. 934-942.

76. Tan, M., et al., *A new approach to develop computer-aided detection schemes of digital mammograms.* Physics in Medicine & Biology, 2015. **60**(11): p. 4413.

77. Justice, M.J. and P. Dhillon, *Using the mouse to model human disease: increasing validity and reproducibility*. 2016, The Company of Biologists Ltd.

78. Hiroshima, Y., et al., *Selective efficacy of zoledronic acid on metastasis in a patient-derived orthotopic xenograph (PDOX) nude-mouse model of human pancreatic cancer.* Journal of surgical oncology, 2015. **111**(3): p. 311-315.

79. Orbay, H., et al., *Developing a Nanoparticle Delivered High-efficacy Treatment for Infantile Hemangiomas Using a Mouse Hemangioendothelioma Model.* Plastic and reconstructive surgery, 2016. **138**(2): p. 410.

80. Zhang, F., et al., *Characterization of drug responses of mini patient-derived xenografts in mice for predicting cancer patient clinical therapeutic response.* Cancer Communications, 2018. **38**(1): p. 60.

81. Pence, B.C., *Calcium and Vitamin D Effects on Tumor Promotion in Rat Colon and Mouse Skin*, in *Calcium, Vitamin D, and Prevention of Colon Cancer*. 2018, CRC Press. p. 191-208.

82. Han, L., et al., *An-te-xiao capsule inhibits tumor growth in non-small cell lung cancer by targeting angiogenesis.* Biomedicine & Pharmacotherapy, 2018. **108**: p. 941-951.

83. Lyons, S.K., *Imaging mouse models of cancer.* The Cancer Journal, 2015. **21**(3): p. 152-164.

84. Oughterson, A., R. Tennant, and E. Lawrence, *The tumor response and stroma reaction following X-ray of a transplantable tumor in inbred strains of mice.* The Yale journal of biology and medicine, 1940. **12**(4): p. 419.

85. Shuhendler, A.J., et al., *Molecular magnetic resonance imaging of tumor response to therapy.* Scientific reports, 2015. **5**(1): p. 1-14.

86. Clark, A.S., M.C. Lynch, and D. Mankoff, *Using nuclear medicine imaging in clinical practice: update on PET to guide treatment of patients with metastatic breast cancer.* Oncology, 2014. **28**(5).

87. Li, X., et al., *Breast osteoblastoma and recurrence after resection: Demonstration by color Doppler ultrasound.* Journal of X-ray Science and Technology, 2017. **25**(5): p. 787-791.

88. Aerts, H.J., et al., *Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.* Nature communications, 2014. **5**(1): p. 1-9.

89. VanOsdol, J., et al., *Sequential HIFU heating and nanobubble encapsulation provide efficient drug penetration from stealth and temperature sensitive liposomes in colon cancer.* Journal of Controlled Release, 2017. **247**: p. 55-63.

90. Aghaei, F., et al., *Applying a new quantitative global breast MRI feature analysis scheme to assess tumor response to chemotherapy.* Journal of Magnetic Resonance Imaging, 2016. **44**(5): p. 1099-1106.

91. Tang, X., *Texture information in run-length matrices.* IEEE transactions on image processing, 1998. **7**(11): p. 1602-1609.

92. Eisenhauer, E.A., et al., *New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1).* European journal of cancer, 2009. **45**(2): p. 228-247.

93. Cohen, J., *Statistical Power Analysis Jbr the Behavioral.* Sciences. Hillsdale (NJ): Lawrence Erlbaum Associates, 1988: p. 18-74.

94. Wang, X., et al., *Improving the performance of computer-aided detection of subtle breast masses using an adaptive cueing method.* Physics in Medicine & Biology, 2012. **57**(2): p. 561.

95. Bokemeyer, C., et al., *Efficacy according to biomarker status of cetuximab plus FOLFOX-4 as first-line treatment for metastatic colorectal cancer: the OPUS study.* Annals of oncology, 2011. **22**(7): p. 1535-1546.

96. Singh, S. and C. Law, *Chromogranin A: a sensitive biomarker for the detection and post-treatment monitoring of gastroenteropancreatic neuroendocrine tumors.* Expert review of gastroenterology & hepatology, 2012. **6**(3): p. 313-334.

97. Janes, H., M.D. Brown, and M.S. Pepe, *Designing a study to evaluate the benefit of a biomarker for selecting patient treatment.* Statistics in medicine, 2015. **34**(27): p. 3503-3515.

98. Zargari, A., et al., *Prediction of chemotherapy response in ovarian cancer patients using a new clustered quantitative image marker.* Physics in Medicine & Biology, 2018. **63**(15): p. 155020.

99. Zheng, B., et al., *Mammography with computer-aided detection: reproducibility assessment—initial experience.* Radiology, 2003. **228**(1): p. 58-62.

100. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: a cancer journal for clinicians, 2018. **68**(6): p. 394-424.

101. Biondi, A., et al., *Neo-adjuvant chemo (radio) therapy in gastric cancer: current status and future perspectives.* World journal of gastrointestinal oncology, 2015. **7**(12): p. 389.

102. Fukagawa, T., et al., *A prospective multi-institutional validity study to evaluate the accuracy of clinical diagnosis of pathological stage III gastric cancer (JCOG1302A).* Gastric Cancer, 2018. **21**(1): p. 68-73.

103. Wang, F.-H., et al., *The Chinese Society of Clinical Oncology (CSCO): clinical guidelines for the diagnosis and treatment of gastric cancer.* Cancer communications, 2019. **39**(1): p. 1-31.

104. Coccolini, F., et al., *Intraperitoneal chemotherapy in advanced gastric cancer. Meta-analysis of randomized trials.* European Journal of Surgical Oncology (EJSO), 2014. **40**(1): p. 12-26.

105. Ishigami, H., et al., *Phase III trial comparing intraperitoneal and intravenous paclitaxel plus S-1 versus cisplatin plus S-1 in patients with gastric cancer with peritoneal metastasis: PHOENIX-GC trial.* Journal of Clinical Oncology, 2018. **36**(19): p. 1922-1929.

106. Sun, Z.-Q., et al., *Radiomics study for differentiating gastric cancer from gastric stromal tumor based on contrast-enhanced CT images.* Journal of X-ray Science and Technology, 2019. **27**(6): p. 1021-1031.

107. Wang, L., et al. *Computer-aided staging of gastric cancer using radiomics signature on computed tomography imaging*. in *Medical Imaging 2020: Computer-Aided Diagnosis*. 2020. International Society for Optics and Photonics.

108. Zheng, B., et al., *Interactive computer-aided diagnosis of breast masses: computerized selection of visually similar image sets from a reference library.* Academic radiology, 2007. **14**(8): p. 917-927.

109. Danala, G., et al., *Classification of breast masses using a computer-aided diagnosis scheme of contrast enhanced digital mammograms.* Annals of biomedical engineering, 2018. **46**(9): p. 1419-1431.

110. Gundreddy, R.R., et al., *Assessment of performance and reproducibility of applying a content-based image retrieval scheme for classification of breast lesions.* Medical physics, 2015. **42**(7): p. 4241-4249.

111. Rajaei, A. and L. Rangarajan, *Wavelet features extraction for medical image classification.* Int. J. Eng. Sci, 2011. **4**: p. 131-141.

112. Hazra, D., *Texture recognition with combined GLCM, wavelet and rotated wavelet features.* International Journal of Computer and Electrical Engineering, 2011. **3**(1): p. 146.

113. Mirniaharikandehei, S., et al., *Developing a quantitative ultrasound image feature analysis scheme to assess tumor treatment efficacy using a mouse model.* Scientific reports, 2019. **9**(1): p. 1-10.

114. Ahmadi, N. and G. Akbarizadeh, *Iris tissue recognition based on GLDM feature extraction and hybrid MLPNN-ICA classifier.* Neural Computing and Applications, 2020. **32**(7): p. 2267-2281.

115. Zhao, F. and C.J. Desilva. *Use of the Laplacian of Gaussian operator in prostate ultrasound image processing*. in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286)*. 1998. IEEE.

116. Wang, Q., et al., *Hierarchical feature selection for random projection.* IEEE transactions on neural networks and learning systems, 2018. **30**(5): p. 1581-1586.

117. Mekhalfi, M.L., et al., *Fast indoor scene description for blind people with multiresolution random projections.* Journal of Visual Communication and Image Representation, 2017. **44**: p. 95-105.

118. Suhaimi, N.F.M. and Z.Z. Htike. *Comparison of Machine Learning Classifiers for dimensionally reduced fMRI data using Random Projection and Principal Component Analysis*. in *2019 7th International Conference on Mechatronics Engineering (ICOM)*. 2019. IEEE.

119. Xie, H., J. Li, and H. Xue, *A survey of dimensionality reduction techniques based on random projection.* arXiv preprint arXiv:1706.04371, 2017.

120. Aggarwal, C.C., A. Hinneburg, and D.A. Keim. *On the surprising behavior of distance metrics in high dimensional space*. in *International conference on database theory*. 2001. Springer.

121. Saunders, C., et al., *Subspace, Latent Structure and Feature Selection: Statistical and Optimization Perspectives Workshop, SLSFS 2005 Bohinj, Slovenia, February 23-25, 2005, Revised Selected Papers*. Vol. 3940. 2006: Springer.

122. Dasgupta, S. and A. Gupta, *An elementary proof of a theorem of Johnson and Lindenstrauss.* Random Structures & Algorithms, 2003. **22**(1): p. 60-65.

123. Pechenizkiy, M., A. Tsymbal, and S. Puuronen. *PCA-based feature transformation for classification: issues in medical diagnostics*. in *Proceedings. 17th IEEE Symposium on Computer-Based Medical Systems*. 2004. IEEE.

124. Peng, H., F. Long, and C. Ding, *Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy.* IEEE Transactions on pattern analysis and machine intelligence, 2005. **27**(8): p. 1226-1238.

125. Zeng, X., et al. *Feature selection using recursive feature elimination for handwritten digit recognition*. in *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. 2009. IEEE.

126. Tibshirani, R., *Regression shrinkage and selection via the lasso.* Journal of the Royal Statistical Society: Series B (Methodological), 1996. **58**(1): p. 267-288.

127. Fernández, A., et al., *SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary.* Journal of artificial intelligence research, 2018. **61**: p. 863-905.

128. Wang, K.-J., et al., *A hybrid classifier combining Borderline-SMOTE with AIRS algorithm for estimating brain metastasis from lung cancer: A case study in Taiwan.* Computer methods and programs in biomedicine, 2015. **119**(2): p. 63-76.

129. Yan, S., et al., *Improving lung cancer prognosis assessment by incorporating synthetic minority oversampling technique and score fusion method.* Medical Physics, 2016. **43**(6Part1): p. 2694-2703.

130. Hu, R., X. Li, and Y. Zhao. *Gradient boosting learning of Hidden Markov models*. in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*. 2006. IEEE.

131. McHugh, M.L., *Interrater reliability: the kappa statistic.* Biochemia medica: Biochemia medica, 2012. **22**(3): p. 276-282.

132. Seevaratnam, R., et al., *How useful is preoperative imaging for tumor, node, metastasis (TNM) staging of gastric cancer? A meta-analysis.* Gastric cancer, 2012. **15**(1): p. 3-18.

133. Liu, S., et al., *CT textural analysis of gastric cancer: correlations with immunohistochemical biomarkers.* Scientific reports, 2018. **8**(1): p. 1-9.

134. Li, R., et al., *Detection of gastric cancer and its histological type based on iodine concentration in spectral CT.* Cancer Imaging, 2018. **18**(1): p. 1-10.

135. Kuhn, M. and K. Johnson, *An introduction to feature selection,* in *Applied predictive modeling*. 2013, Springer. p. 487-519.

136. Tan, M., J. Pu, and B. Zheng, *Optimization of breast mass classification using sequential forward floating selection (SFFS) and a support vector machine (SVM) model.* International journal of computer assisted radiology and surgery, 2014. **9**(6): p. 1005-1020.

137. Khalid, S., T. Khalil, and S. Nasreen. *A survey of feature selection and feature extraction techniques in machine learning*. in *2014 Science and Information Conference*. 2014. IEEE.

138. Chandrashekar, G. and F. Sahin, *A survey on feature selection methods.* Computers & Electrical Engineering, 2014. **40**(1): p. 16-28.

139. Mirniaharikandehei, S., et al., *Applying a random projection algorithm to optimize machine learning model for predicting peritoneal metastasis in gastric cancer patients using CT images.* Computer Methods and Programs in Biomedicine: p. 105937.

140. Blagus, R. and L. Lusa, *SMOTE for high-dimensional class-imbalanced data.* BMC bioinformatics, 2013. **14**: p. 106-106.

141. Wang, T., et al., *Correlation between CT based radiomics features and gene expression data in non-small cell lung cancer.* Journal of X-ray Science and Technology, 2019. **27**(5): p. 773-803.

142. Galloway, M.M., *Texture analysis using grey level run lengths.* NASA STI/Recon Technical Report N, 1974. **75**: p. 18555.