INFORMS TutORials in Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Quantitative Imaging System for Cancer Diagnosis and Treatment Planning: An Interdisciplinary Approach

Teresa Wu, Nathan Gaw, Yanzhe Xu, Jing Li, Lujia Wang, Yinlin Fu, Alvin Silva, Christine Zwart, Mitesh Borad, Thomas DeLeon, Bhavika Patel

To cite this entry: Teresa Wu, Nathan Gaw, Yanzhe Xu, Jing Li, Lujia Wang, Yinlin Fu, Alvin Silva, Christine Zwart, Mitesh Borad, Thomas DeLeon, Bhavika Patel. Quantitative Imaging System for Cancer Diagnosis and Treatment Planning: An Interdisciplinary Approach. *In* INFORMS TutORials in Operations Research. Published online: 03 Oct 2017; 153-177.
https://doi.org/10.1287/educ.2017.0173

TUTORIALS IN
OPERATIONS RESEARCH
INFORMS 2017

 | ISBN 978-0-9906153-0-9
https://doi.org/10.1287/educ.2017.0173

Quantitative Imaging System for Cancer Diagnosis and Treatment Planning: An Interdisciplinary Approach

Teresa Wu, Nathan Gaw, Yanzhe Xu, Jing Li, Lujia Wang, Yinlin Fu
School of Computing, Informatics, and Decision Systems Engineering, Ira Fulton Schools of Engineering, Arizona State University, Tempe, Arizona 85281
{teresa.wu@asu.edu, ngaw1@asu.edu, yanzhe.xu@asu.edu, jing.li.8@asu.edu, lujia.wang.1@asu.edu, yinlinfu@asu.edu}

Alvin Silva, Christine Zwart, Bhavika Patel
Department of Radiology, Mayo Clinic, Scottsdale, Arizona 85259
{silva.alvin@mayo.edu, zwart.christine@mayo.edu, patel.bhavika@mayo.edu}

Mitesh Borad, Thomas DeLeon
Department of Hematology and Oncology, Mayo Clinic, Scottsdale, Arizona 85259
{borad.mitesh@mayo.edu, deleon.thomas@mayo.edu}

Abstract During the past decade, with breakthroughs in systems biology, precision medicine has emerged as a novel paradigm that has transformed healthcare. Precision medicine is an approach for disease treatment and prevention that takes into account individual variability where medical imaging is a key component. This tutorial focuses on research investigating the roles of medical imaging in cancer diagnosis and treatment planning. While the cornerstone of the imaging research is mathematical and statistical modeling, the research has to take a multidisciplinary and systematic approach because of the nature of the problem. We offer a comprehensive review and discussion on four important components that form an imaging pipeline: imaging preprocessing, imaging feature extraction, feature dimensionality reduction, and classification. To illustrate the clinical relevance, our in-house-developed system, the imaging multitexture disease diagnosis system, is presented with two clinical case studies: one on breast cancer diagnosis using contrast-enhanced digital mammography imaging and the other on cholangiocarcinoma using computed tomography imaging. The future directions of the imaging research are highlighted in the end.

Keywords clinical decision support; multitexture data integration; disease diagnosis; cancer; treatment planning

1. Introduction

Cancer is the leading cause of death for Americans aged 40–79 years (Sagon [50]). According to the National Cancer Institute [44], nearly 1.7 million new cancer cases and 595,000 cancer deaths (one out of every four deaths) were estimated to occur in the United States in 2016. In fact, approximately 40% of men and women will be diagnosed with cancer at some point during their lifetimes. National expenditures for cancer care in the United States totaled nearly $125 billion in 2010 and could reach $156 billion in 2020 (see National Cancer Institute [44] for details). In clinical cancer practice, tissue biopsy is the gold standard because of its accuracy in diagnosis and treatment response evaluation. Although valuable, tissue biopsies have several limitations: they are invasive, are uncomfortable, and expose the patient to significant risk (Field et al. [17]). In addition, a biopsy sample is usually done

for a single location of a tumor at a single time point. This fails to capture the spatial genomic diversity of the tumor and its temporal changes, which have been recognized to account for diagnostic and treatment failure of a number of aggressive cancers. Imaging, on the other hand, provides macroscopic information on tumor morphology, physiology, and function (Zinn et al. [72]). It is efficient and noninvasive; thus imaging is naturally performed for the entire tumor and at multiple time points in follow-ups. In recent years, cutting-edge imaging technologies have been developed with high-quality, high-resolution images to provide complementary views of the structural, functional, and mechanical properties of a tumor. This has pushed imaging to an unprecedented opportunity for cancer care toward personalized and precision medicine. Consequently, there has been considerable research over the last decade aimed at screening imaging features to identify patterns for tumor phenotype prediction (see Davnall et al. [12], Kassner and Thornhill [29], Skogen et al. [55], Zacharaki et al. [68]). In general, mathematical and statistical models are developed to assist the prediction, but their usefulness depends on the imaging features investigated. To enhance the performance of these models, research has been devoted to developing methods that extract texture features from medical images.

In medical images, texture characterization has become an important problem. Texture can be thought of as the local characteristic pattern of image intensity that identifies a tissue. Since changes in the local texture will cause changes in the local spatial frequency, texture analysis evaluates the local spectral or frequency contents to extract multiple features on the pixel/voxel level. The analysis of these texture features provides a meaningful way to assist in disease diagnosis and staging. Texture analysis is an ongoing field of research, with broad applications that include the segmentation of specific anatomical structures, the detection of lesions, the distinction between pathological and healthy tissue in different organs, and the associations with genomic features. Extensive research has investigated various texture analysis methods for medical image processing, which can be generally classified into four categories: structural methods, model-based methods, statistical methods, and transformation-based methods. In structural methods, texture is characterized by feature primitives and their spatial arrangements. One example of the primitive is a square object that is represented in terms of the straight lines that form its border. While simple, structural texture analysis techniques may be limited in practical use since they can only describe regular objects such as a square or sphere. Model-based methods represent texture via the use of generative image models (e.g., fractals). The main idea behind a fractal model is the property of self-similarities—that is, an object can be decomposed into smaller similar copies of itself. A metric called fractal dimension is introduced to describe the disorder of an object—i.e, the higher the dimension value, the more complicated the object. One criticism of the model-based approach is the computational complexity involved in the estimation of the large number of model parameters. Statistical methods identify the texture by considering the spatial distribution of intensity values at each pixel/voxel in the images and compute a set of second-order (or even higher-order) statistics from the distributions of the local features. Examples of these methods are gray level co-occurrence matrix (GLCM) and local binary patterns (LBPs). Transform-based method converts the images into new forms using spatial frequency properties of the pixel/voxel intensity variation. For example, wavelet transform analyzes the frequency content of an image within different scales and frequency directions. Wavelet coefficients corresponding to the scales and directions can be derived to describe the image properties as texture features. Some examples of transform-based methods are Gabor wavelet transform, Fourier transform, and S-transform.

Since each type of texture feature reveals different aspects of the image, using a multitextural approach has become an emerging research trend. The general multitextural approach includes four steps. First, regions of interest (ROIs) are placed over the tissue area. To address the intratumor and intertumor heterogeneity, ROI normalization is implemented to bring the imaging intensity scales to the same level. Next, some commonly studied texture

algorithms such as GLCMs, LBPs, and Gabor filters are parameterized based on the imaging modality and disease type, and are applied to the normalized ROIs. Given that imaging features are typically high dimensional, and some features are naturally highly correlated because of their spatial proximity or functional similarity, dimension reduction techniques such as principal component analysis (PCA) are applied as the third step. Finally, multiparametric predictive models are developed based on the principle components (PCs) to assist with disease diagnosis.

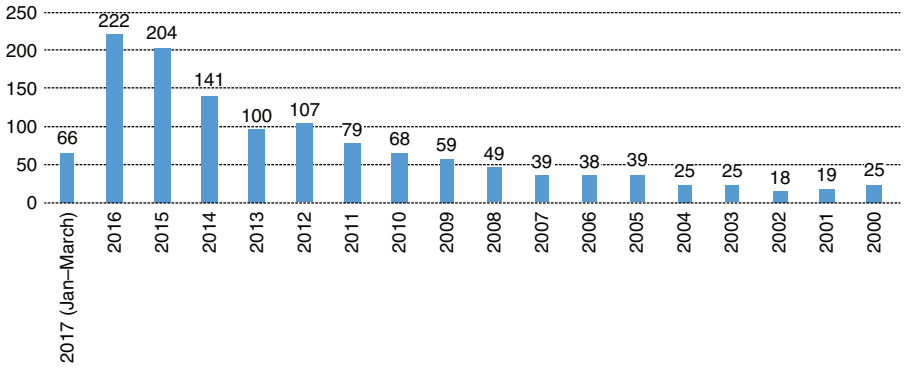
Following the same workflow, we develop a multitexture pipeline, the imaging multitexture disease diagnosis system, termed iMT-DDS. In this tutorial, we present the formulation of iMT-DDS and demonstrate its clinical applications. Specifically, two case studies are conducted. In the first study, we explore the applicability of iMT-DDS to breast cancer diagnosis using contrast-enhanced digital mammography (CEDM) imaging. The data set includes 99 patients (21 benign and 78 malignant). The best-performing model from leave-one-out cross validation (LOOCV) gives 91% sensitivity, 71% specificity, and 86% overall accuracy. In the second study, we apply iMT-DDS to discover the associations between imaging features with genome types (a more challenging problem in the area of radiogenomics). We study cholangiocarcinoma using computed tomography (CT) imaging. The data set includes 33 patients (15 fibroblast growth factor receptor 2 positive (FGFR2+) and 18 FGFR2 negative (FGFR2−)). The LOOCV classification performance gives 87% sensitivity, 94% specificity, and 91% overall accuracy.

The rest of the paper is structured as follows: Section 2 provides a general overview of different imaging modality and imaging research for cancer. Section 3 gives a detailed discussion of imaging modalities and an imaging pipeline that consists of imaging preprocessing, imaging feature extraction, feature selection, and classification. In Section 4, two clinical case studies on breast cancer and cholangiocarcinoma, using our in-house-developed imaging pipeline, iMT-DDS, are presented. The conclusion is then drawn in Section 5, followed by a discussion on future directions.

2. Overview of Imaging Research for Cancer

Modern imaging technologies allow for visualization of multidimensional and multiparameter data. Imaging is increasingly used to measure physical parameters such as tissue properties and to glean temporal insight on biological function. CT, also commonly referred to as a CAT scan, is an imaging technique that combines multiple x-ray projections taken from different angles to produce detailed cross-sectional views of the body part of interest. CT has relatively low soft tissue contrast for tumor and surrounding tissue. But with iodinated contrast agents, organs and tumors can also be detected. Given CT's fast imaging time and high spatial resolution, it has been used to image lung tumors and bone metastasis. magnetic resonance imaging (MRI) is performed by placing a subject in a strong magnetic field, typically 1.5 or 3 tesla for human scanners, which aligns the hydrogen nuclei spins in a direction parallel to the field. Like CT, MRI traditionally creates a two-dimensional (2D) image of a thin "slice" of the body and is therefore considered a tomographic imaging technique. These days, advancements in magnetic resonance (MR) techniques enable a three-dimensional (3D) view of the images. Since there is no use of ionizing radiation, MR has no health concern of dose exposure, as opposed to CT. MRI has proven to be highly effective in diagnosing a number of conditions by showing the difference between normal and diseased soft tissues of the body such as breast, heart, and kidney. Positron emission tomography (PET) is a nuclear imaging technique that incorporates radioactive tracers in the image acquisition. It provides physicians with information about how tissues and organs function and is therefore called a functional imaging technique. By using radioactive tracers, three-dimensional images can be reconstructed to show the concentration and locations of metabolic molecules of interest. PET is ideally suited for monitoring molecular events early in the course of a disease, as well

FIGURE 1. Number of publications about cancer imaging over 2000–2017 (from PubMed).



as during pharmacological or radiation therapy. Digital mammography (DM) is a specialized modality for breast imaging. It uses low-dose x-rays to detect breast cancer early and has been adopted as the technique in the breast cancer screening program.

As reviewed, each imaging modality has its own unique value to clinical applications, and they all present the images in the form of two or three dimensions to the radiologists for cancer diagnosis and staging. However, subtle heterogeneity of tissues and lesions may not be perceptible to the human eye, even for experienced radiologists. Texture analysis, a computational technique that quantifies the intrinsic nuances in the images, may address this challenge. Texture features derived from imaging provide a quantitative analysis that complements the molecular approaches for cancer detection, diagnosis, and prognosis (see Angell et al. [1], Dunn et al. [15], Laurinavicius et al. [34, 35]). Diagnostic and prognostic values of texture analysis have been experimentally established in a large number of studies across different imaging modalities including CT, PET, MRI, and others. In this tutorial, to provide a comprehensive review of the state-of-the-art research in this field, we first use the keywords "texture," "imaging," and "cancer" to search the publications from this arena through PubMed (https://www.ncbi.nlm.nih.gov/pubmed/). Since a new term, "radiomics," emerges in 2012, we also add "radiomics" to the search. As shown in Figure 1, the number of publications has grown exponentially over the years. Even in the first quarter of the year 2017, there have been 66 publications indicating a growing interest in this field. For the remainder of this section, we provide a thorough review of clinical studies that employ GLCM, LBP, and Gabor textures for classification.

2.1. Modeling with GLCM Features

One common application of GLCM is in CT imaging. Lo et al. [37] study 8 intensity-based histogram features and 34 GLCM texture features to investigate the effects of dose level and reconstruction method on density and texture features computed from CT lung nodules. It is concluded that histogram mean is the most robust feature, followed by the summation of products between intensities and probabilities, and finally the GLCM features, which vary widely. Pham et al. [46] implement two complementary methods of texture analysis, GLCM and experimental semivariogram, on CT imaging aiming to improve predictive power of evaluating mediastinal lymph node in lung cancer. A sensitivity of 75% and specificity of 90% are achieved using a logistic regression model. In another lung cancer study using CT imaging, Wu et al. [64] extract tumor shape, size, and intensity statistics, as well as texture features from GLCM and gray level run-length matrices. They conclude multiple imaging features are significantly associated with tumor histology. The authors also employ multiple classifiers such as random forests, naive Bayes, and k-nearest neighbors to develop multivariate predictive models. The best-performing classifier is naive Bayes. Hanania et al. [22]

investigate 360 imaging features including the intensity, shape, and textures of CT images to evaluate malignant potentials of intraductal papillary mucinous neoplasms. The best classifier is the logistic regression model. Another interesting study by Yoon et al. [67] extracts histogram features (kurtosis and skewness) and GLCM from CT images of human epidermal growth factor receptor 2 (HER2)-positive patients to identify the association between the imaging features with the survival rates after trastuzumab treatment.

GLCM has also been used with sonogram imaging. Song et al. [56] evaluate the use of GLCM texture features from sonogram images to predict thyroid nodes. Six classifiers—support vector machine (SVM), random tree, random forest, boost, logistic, and artificial neural network—are developed. The logistic model shows the best performance with 79% sensitivity and 79% specificity using 10-fold cross validation. Lashkari [33] studies the texture features from infrared breast thermography for early detection of breast cancer. Similarly, with 23 features including statistical, morphological, frequency domain, histogram, and GLCM texture, multiple feature selection algorithms are explored, including minimum redundancy and maximum relevance, sequential forward selection, sequential backward selection, and genetic algorithms. Based on the selected features, again, different classifiers are implemented, such as AdaBoost, SVM, k-nearest neighbors, naive Bayes, and probabilistic neural network.

PET imaging has been the focus of some studies since it captures the physiological information of tumors. Using phantom data, Wang et al. [60] compare the performance of the H index, a feature defined from the PET SUV value (imaging-specific feature) and GLCM features. Interestingly enough, the authors conclude that the H index indeed outperforms GLCM in characterizing SUV heterogeneities. Doumou et al. [14] study the tumor heterogeneities using PET imaging by exploring multiple texture algorithms including GLCM, the gray level run length matrix, the neighborhood gray-tone difference matrix, the gray level size zone matrix, and the fractal analysis method. They conclude that heterogeneity measurement precision in PET is largely influenced by the image process and the texture algorithms.

Chaddah et al. [7] study the applicability of multiple texture algorithms on multispectral pathological images. After evaluating the Laplacian of the Gaussian (LoG), discrete wavelets (DWs), and GLCM, the authors conclude that GLCM texture features outperform LoG and DW. Yet the higher performance is achieved by combining all texture features. Zhang et al. [71] study diffusion-weighted images and the apparent diffusion coefficient map to distinguish between low-grade and high-grade bladder cancer. Specifically, histogram and GLCM features are extracted on MR images that are then used to develop an SVM model embedded with recursive feature selection strategy. Chen et al. [9], on the other hand, study the use of GLCM on conventional MR methods—T1-weighted, T2-weighted, fluid-attenuated inversion recovery, and contrast-enhanced T1 imaging—for glioblastoma. Five GLCM texture features are extracted for each of the four MR protocols, and the best model was found to achieve 75% sensitivity and 100% specificity.

2.2. Modeling with LBP Features

Reyad et al. [49] extract both statistical features and LBP features from each region taken from mammogram images. SVM classifiers are developed based on these two sets of features, and similar accuracy (98%) is achieved. A marginally improved accuracy of 99% is achieved by taking both statistical and LBP features. Gangeh et al. [18] study the LBP features and intensity-based features from ultrasound images to assess the breast cancer treatment efficacy. Cai et al. [6] develop a phase congruency-based binary pattern, which is a combination of phase congruency with LBP features. These novel integrated features are then used for a SVM classifier to diagnose breast cancer, and a high 0.894 area under the receiver operating characteristic curve (AUC) is achieved.

2.3. Modeling with Gabor Features

Lu et al. [39] study various features including shape features and Gabor filters from MR images for breast cancer diagnosis and achieve an AUC of 0.962. Suganthi and Ramakrishnan [57] attempt to distinguish normal and abnormal tissue in breast thermal images using Gabor wavelet transform. In a separate study, Han et al. [21] study the pulmonary nodules from CT images. Three texture analysis algorithms, GLCM, Gabor, and LBP from 2D images to 3D space, are studied. GLCM textures achieve the highest AUC value of 0.927. Similarly, Atupelage et al. [3] develop a bag-of-feature approach that incorporates Gabor, LBP, and GLCM features from histopathological images for hepatocellular carcinoma grading.

2.4. Summary

Given the importance of imaging in cancer diagnosis, staging, and treatment assessment, enormous research efforts have been devoted to extract meaningful imaging features from which mathematical and statistical prediction models are developed. From this review, we can draw three observations. First, texture analysis is a generalized technique with broad applications in different imaging modalities. Second, features from different sources, such as morphology, statistical, and texture, are of interest in most research. In the category of texture, multiple texture algorithms have been jointly evaluated in a large number of studies. Third, given the large number of features, some research has employed multiple feature selection and multiple predictive modeling in an attempt to achieve the best possible predictive power. This is probably due to the fundamental challenges in these types of studies—that is, disease type differs, individual patient differs, and imaging modality differs. As a result, to fully take advantage of a scientific data-driven approach, the study design should be inclusive. In other words, the study should consider multiple types of features such as intensity histogram and different textures. Additionally, as the "no free lunch" theorem by Wolpert and Macready [63] states, no algorithm can outperform any other algorithm when performance is amortized over all functions. So the study design should include multiple predictive modeling methods as well.

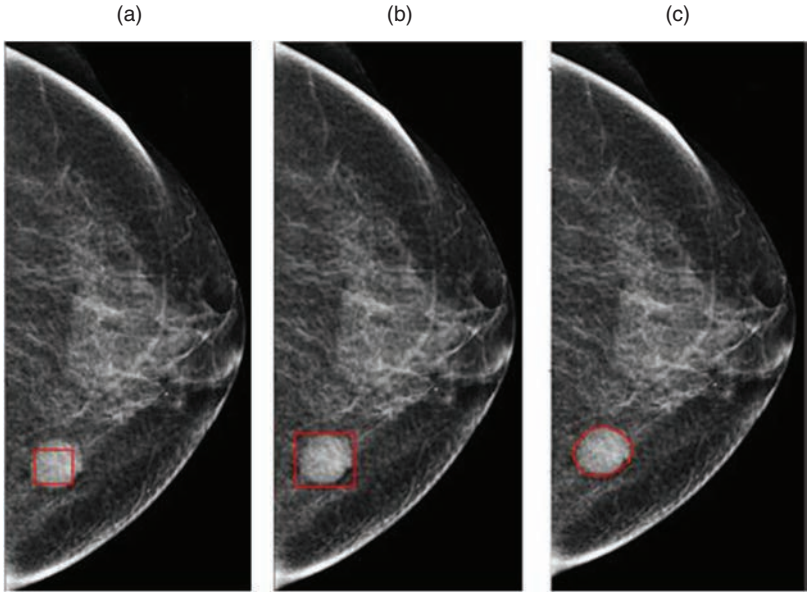
In the next section, a four-stage imaging pipeline for cancer research is presented. This pipeline starts with imaging preprocessing, followed by the imaging feature extraction, feature selection, and classification. This pipeline is essentially the core of quantitative imaging, the emerging interdisciplinary research spanning medicine, engineering, computer science, and mathematics.

3. Quantitative Imaging System for Disease Diagnosis

A quantitative imaging system involves a process that consists of five steps: imaging acquisition, ROI definition/preprocessing, feature extraction, feature selection, and classification (see Figure 2). For the audience of this tutorial, we focus on the last four steps. Note that none of these steps is specific, and the methods have to be chosen according to the application. The texture outcome can be considerably affected depending on the methodology used throughout the process.

Figure 2. Workflow of a quantitative imaging system.

FIGURE 3. Approaches for defining a ROI over a breast tumor: (a) a small square inside the tumor, (b) a big square outside the tumor, and (c) delineation of the tumor.



3.1. ROI Definition and Preprocessing

Initially, texture features are extracted from a predefined section of a 2D or 3D image. The region of interest (for 2D images) or volume of interest (for 3D images) is usually placed over the tissue area. Since manual segmentation of ROIs outperforms automatic methods, it is still considered the gold standard clinically (Harrison et al. [24], Loizou et al. [38], Sanz-Cortez et al. [51], Shi et al. [53]). In general, there are three approaches to define these ROIs. As seen in Figure 3(a), we can place a square inside the area of tissue to be analyzed. While only information of the tissue of interest is captured, some texture details may be lost because this ROI does not cover the entire area. Alternatively, we can position a smallest bounding box enclosing the tissue area; see Figure 3(b). But this method may introduce noise by including adjacent image pieces outside the tissue area. The third approach, shown in Figure 3(c), delineates the entire tissue area or lesion. Although it may be a better approach, since the whole area of interest is included, it is time consuming and labor intensive. In addition, it requires precise delineation to achieve the desired outcomes. In comparison, the first two approaches are relatively easy to implement because of their use of common shapes. Note that a potential issue facing all three approaches is border effect—that is, the neighborhood of border pixels will include pixels outside the ROI when the texture features are derived. If the ROI is large enough, the border effect may be negligible; otherwise, this issue needs to be addressed to capture accurate texture information of the ROI.

Before applying texture algorithms on the defined ROIs, a necessary preprocessing step is normalization. Some imaging features are not only dependent on local textures but also dependent on ROI properties such as mean intensity and variance (Materka et al. [41]). A number of clinical factors lead to the variations of these ROI properties. For instance, to assess the treatment responses, one patient will have multiple follow-ups with imaging. The imaging intensity is not homogeneously distributed over the course of multiple follow-ups. This variation may be amplified over different patients because of the patient physiological differences even if the imaging acquisition protocol is standardized. Collewet et al. [11] study the effects of ROI normalization on MR images and recommend the $\pm 3\sigma$ approach where the image intensities are normalized between $\mu \pm 3\sigma$, μ is the mean value of gray

levels inside the ROI, and σ is the standard deviation, so that gray levels located outside the range $[\mu - 3\sigma, \mu + 3\sigma]$ are not considered for further analysis. For simplicity, researchers have also employed the min-max approach—that is, using the minimum and maximum intensity values from the ROIs to normalize the intensity value of the ROI pixels. Depending on the imaging modality, the minimum and maximum values may be taken from different ROIs. Taking dynamic MR as an example, normalization is recommended to look into at least two time points of the imaging, the first time point and the second time point. The normalization is conducted over the ROIs from the two time points. The minimum and maximum values of both ROIs are used for normalization (Michoux et al. [42]). To address the patient physiological differences, researchers have also developed an approach that obtains two ROIs from each patient, one from the tumor area and another from the normal tissue area, and the overall minimum and maximum intensity values are used for the normalization.

3.2. Texture Feature Extraction

Feature extraction is the main step in the texture analysis process and involves the computation of texture features from predefined and normalized ROIs. Many algorithms have been proposed to quantify the textures of an image allowing the computation of numerous features. One observation from our literature review is that many studies implement multiple texture feature algorithms with some of the main algorithms of interest being GLCM, LBP, Gabor filters, and their extensions.

3.2.1. GLCM. In GLCM, the co-occurrence matrix is used to extract statistical information about the distribution of pixel pairs in the image of a specific spatial relationship given the direction (e.g., 45°, 90°) and the distance (e.g., 1-pixel separation). The column and row number of GLCM are assigned to specific values, and the element shows the co-occurrence of two corresponding values. Once the co-occurrence matrix is derived, 13 texture features are calculated: angular second moment, contrast, correlation, sum of squares (variance), inverse difference moment, sum of average, sum of variance, sum of entropy, entropy, difference variance, difference entropy, and two types of information measures of correlation (Haralick and Shanmugam [23]). Since the GLCM features are centrosymmetric in terms of the directions, it is recommended that only directions less than 180° from the positive x axis of the image are used. The features from reversed direction will be covered as a result of centrosymmetry. Four directions (0°, 45°, 90°, and 135°) are typically used in computing the features. The pixel distance is chosen based on the application: a larger distance will allow the detection of coarse areas, while a smaller distance is good for more granulized detections.

Given the separation distance and direction, a co-occurrence matrix can be derived. An example of the co-occurrence matrix is shown in Figure 4. For all the co-occurrence matrices,

FIGURE 4. Example of computing a co-occurrence matrix in GLCM: (a) gray level images (three levels: black, dark gray, and light gray), (b) gray level representation of the image (1 = black, 2: dark gray, 3: light gray), and (c) for 0°, one-pixel separation, two occurrences of 1-2.

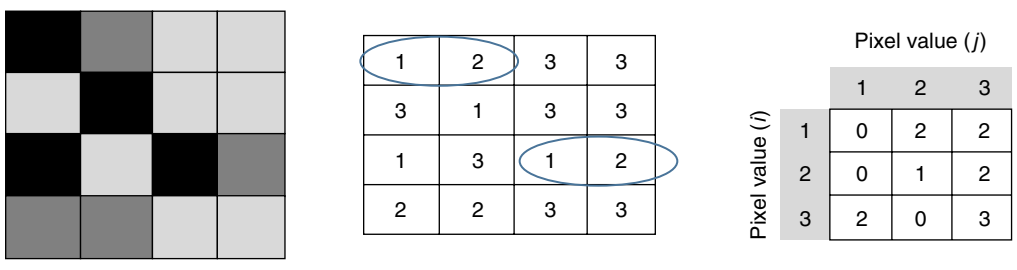
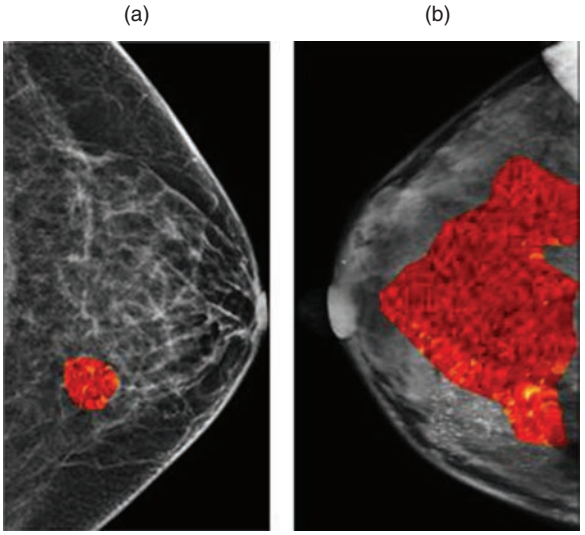


FIGURE 5. Demonstration of GLCM difference variance average feature from contrast-enhanced digital mammography of breast cancer: (a) benign case, (b) malignant case.



the numbers of occurrences between each pixel pair are calculated. Taking grayscale images as example, a black pixel has value 0 and a white pixel has value 255. The number of 0-255 (black-white) pairs across all co-occurrence matrices is derived.

From the co-occurrence matrix, 13 statistical features from these counts are calculated (see Appendix A for the details of the features). As an illustration, we generate a feature map on one of the GLCM texture features (difference variance average feature) from CEDM breast imaging. Figure 5(a) shows the feature map on a benign case, and Figure 5(b) shows the feature map on a malignant case.

3.2.2. LBP. LBP is a texture descriptor introduced by Ojala et al. [45]. It became popular because of its simplicity and good discriminative power. For each pixel within an image, the LBP algorithm compares the gray level of its neighborhood pixels. If the neighborhood pixel has a greater value, 1 is assigned; otherwise, 0 is assigned. After labeling all the neighborhood pixels, an LBP vector consisting of binary values is generated. An example of LBP on a 3×3 patch is shown in Figure 6. As an illustration, Figure 7 shows the feature map on a benign case (Figure 7(a)) and a malignant case (Figure 7(b)). For this 3×3 patch, the LBP vector will have eight digits. A histogram on 256 bins can be generated to describe the image properties. But this number of bins can be adjusted based on the imaging modality and disease type. If the bin size is too small, the size of each bin will be large, which may lose some granularity of the estimated distribution; on the other hand, if the bin size is too large, computational complexity issues arise. This trade-off leads to uniform LBP bin size, which is usually determined from a series of empirical experiments. As a guideline, most clinical research uses 10–20 bins, depending on the data set as well as the computational power of the systems used to run the algorithms. The patch size can also be extended. If the neighborhood size is 2, the surrounding pixels will be 24 pixels. Thus, the number of bins and the bin sizes should be adjusted accordingly.

3.2.3. Gabor Filter Bank. The Gabor filter is a method based on image transformation that produces an image in a space whose coordinate system is related to texture characteristics, such as frequency content or spatial resolution. Commonly studied transform-based methods are Gabor filters, Fourier, and S-transform, among which Gabor filters are known to provide better spatial localization (Larroza et al. [32]). The Gabor filter is a band pass

defined by a frequency and an orientation parameter. In particular, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal wave. Since the Gabor filter shows optimal localization properties in both spatial and frequency domains, it has proven its advantages in edge detection and texture segmentation problems (Ratha et al. [48]). However, their usefulness may be limited because there is no single filter resolution at which a spatial structure can be localized (Materka [40]). As a result, a Gabor filter bank, a set of predefined Gabor filters with the combination of frequencies and orientations of the sinusoidal wave and the variances of the Gaussian kernels, is often constructed. The filtered images with different Gabor filters are then convolved, and the mean value and standard deviation of each filtered image are calculated as texture features. For a sinusoidal wave in a 2D space, the direction parameter describes the angle of the wave function; it is easy to see that this parameter is also centrosymmetric on the original directions. Thus four directions are commonly used in Gabor filters: 0°, 45°, 90°, and 135°. The frequency parameter decides the location of a series of peaks on the sinusoidal wave. These peaks are the band-pass position on the wave and should be selected based on the size of the image as well as the potential textures embedded in the image. For the Gaussian kernel function in Gabor filters, the variance also affects the width of the band pass. These two parameters are both related to the image properties. In Figure 8, we show an example of a feature map on Gabor filters with the standard deviation being 0.6 and frequency being 0.3.

As seen in Figures 5, 7, and 8, the color scales may reflect the subtle differences between the benign case and the malignant case. Joining all the features together will enable the

FIGURE 6. Illustration of LBP algorithm on 3×3 patch, where center pixel x has a vector of 8 bits [1 1 1 0 1 0 0 1 0] and the value is 87.

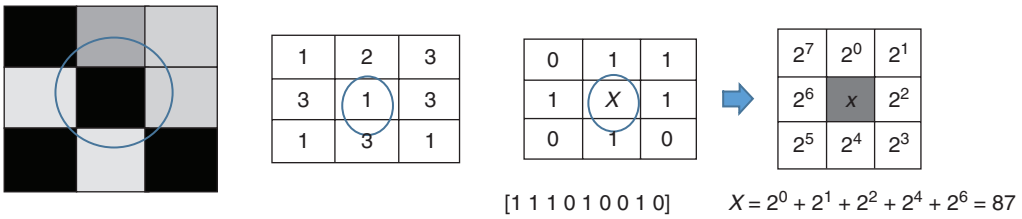


FIGURE 7. Demonstration of LBP feature from contrast-enhanced digital mammography of breast cancer: (a) benign case, (b) malignant case.

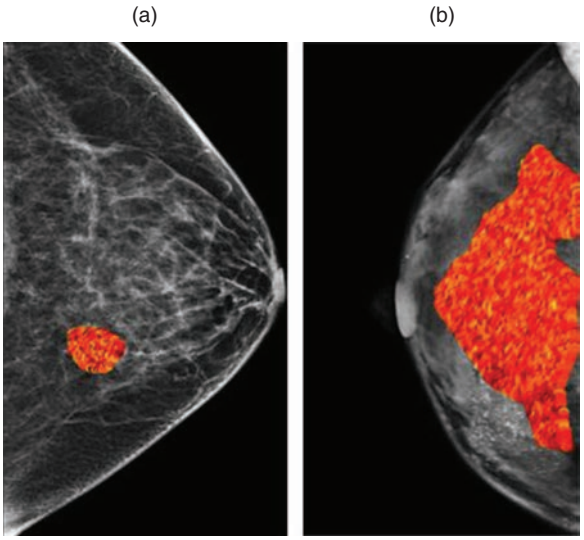
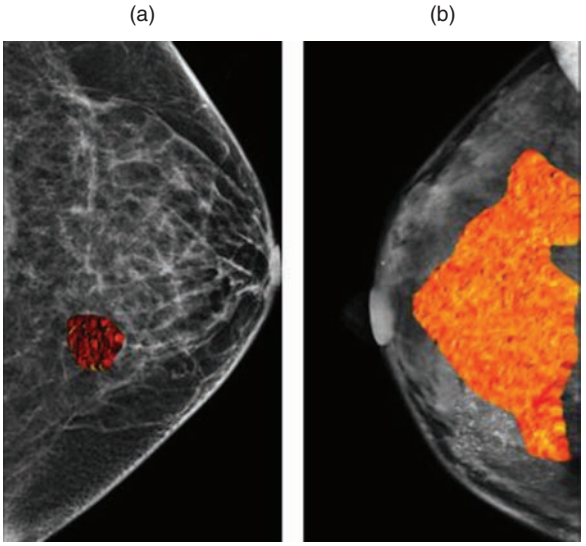


FIGURE 8. Demonstration of Gabor standard deviation (sigma = 0.6, frequency = 0.3) from contrast-enhanced digital mammography of breast cancer: (a) benign case, (b) malignant case.



development of a predictive model that will have discriminative power for disease diagnosis. However, these features may be highly correlated as a result of spatial adjacencies and functional similarity. To reduce the dimensionality and decorrelate the features, some dimension reduction techniques are applied such as principle component analysis (PCA).

3.3. Principle Component Analysis for Feature Dimension Reduction

PCA is a machine learning method that transforms the data into principal components (PCs), linear combinations of the original features (Johnson and Wichern [28]). Using this transformation, the original data set, which may involve many features with high correlations, can often be interpreted by just a few uncorrelated PCs. The first task in PCA is to identify a new coordinate axis that has the largest possible variance in the dataspace. This forms the base axis for the first PC. Following a similar procedure, the feature that is orthogonal to the first PC and has the largest possible variance is used to identify the next direction, the second PC. Iteratively, this procedure continues until the original feature space is transformed to the orthogonal PC space.

3.4. Predictive Modeling

The PCs produced from the previous step are then used to build a classification model. There are many types of classification models that can be chosen. Popular classification models include linear discriminant analysis (LDA; Hu et al. [26], Huang et al. [27]), quadratic discriminant analysis (QDA; Chong et al. [10], Schwedt et al. [52], Zhang et al. [70]), and SVMs (Fan et al. [16], Yang et al. [66], Zhang et al. [69]).

LDA creates a model with a linear classification boundary that optimally divides instances in a binary classification problem. Given $Y \in \{0,1\}$, LDA takes the following form:

$$\log \frac{p(Y=1 \mid \mathbf{z})}{p(Y=0 \mid \mathbf{z})} = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}^{-1} \mathbf{z} - \frac{1}{2}\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 + \frac{1}{2}\boldsymbol{\mu}_0^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_0 + \log \frac{\pi}{1-\pi}, \tag{1}$$

where $\mathbf{z}$ represents the set of PCs, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_0$ are the means of $\mathbf{z}$ for the two classes, $\boldsymbol{\Sigma}$ is the pooled covariance matrix for the two classes, and $\pi = P(Y = 1)$. The classification rule of

LDA is that if $\log(p(Y=1\mid \mathbf{z})/(p(Y=0\mid \mathbf{z}))>0$, assign the sample to class 1 and to class 0 otherwise. Note that $\boldsymbol{\mu}_0$, $\boldsymbol{\mu}_1$, $\boldsymbol{\Sigma}$, and π can be estimated from training data by maximum likelihood estimation.

QDA is similar in formulation to LDA, except that it has a quadratic decision boundary. Given $Y\in\{0,1\}$, QDA takes the following form:

$$\log\frac{p(Y=1\mid \mathbf{z})}{p(Y=0\mid \mathbf{z})}=-\frac{1}{2}\mathbf{z}^T(\boldsymbol{\Sigma}_1^{-1}-\boldsymbol{\Sigma}_0^{-1})\mathbf{z}+(\boldsymbol{\mu}_1^T\boldsymbol{\Sigma}_1^{-1}-\boldsymbol{\mu}_0^T\boldsymbol{\Sigma}_0^{-1})^T\mathbf{z}-\frac{1}{2}\boldsymbol{\mu}_1^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 \tag{2}$$

$$+\frac{1}{2}\boldsymbol{\mu}_0^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0+\log\frac{\pi}{1-\pi}+\log\sqrt{|\boldsymbol{\Sigma}_0|/|\boldsymbol{\Sigma}_1|}, \tag{3}$$

where $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_0$ are the covariance matrices of the two classes (unlike LDA, QDA assumes each class has a unique covariance matrix).

The SVM model has a linear decision boundary and takes the following form:

$$f(\mathbf{z})=\mathbf{s}^T\mathbf{z}+s_0, \tag{4}$$

where $\mathbf{s}$ and s_0 are estimated from the objective function $\min_{\{\mathbf{s},\,s_0,\,\xi\}}\frac{1}{2}\mathbf{s}^T\mathbf{s}+C\boldsymbol{\Sigma}_i\xi_i$ subject to $y_if(\mathbf{z}_i)\geq 1-\xi_i$ and $\xi_i\geq 0\ \forall i$, where C is the penalty parameter, ξ_i is the slack variable for sample i in a training data set, y_i is the class of sample i, and $f(\mathbf{z}_i)$ is the predicted value of sample i.

To select important features that build an accurate model, a wrapper method is employed. The wrapper method is useful for developing an accurate model since it directly considers the accuracy of the classifier in model development (Guyon and Elisseef [20]). Note that wrapper method development itself is a growing research field. Because of space limitations, we only briefly discuss two wrappers used in our studies. One is sequential forward selection (SFS), which can be used to create a simple and accurate model (Whitney [62]). SFS starts with an empty set of features and, in each step, adds a new PC such that the classification accuracy is maximized. SFS stops when it cannot find a PC that will improve the accuracy by a certain threshold chosen by the user (for example, stop adding PCs to the model when the accuracy does not improve by >1%). The second method is the particle swarm optimizer (PSO), a swarm intelligence algorithm that can be embedded in the predictive modeling process for feature selection.

Once an optimal classifier with the selected features is found, the PCs in the models can be reverse transformed to show the contribution of the original features (Gaw et al. [19]). As a result, the contributing features can be identified for clinical interpretations. Interested readers are referred to Hu et al. [25, 26] and Ramkumar et al. [47] for examples.

4. Imaging Multitexture Disease Diagnosis System and Its Clinical Applications

Following the same workflow of the quantitative imaging system laid out in Section 3, we develop iMT-DDS, an imaging multitexture disease diagnosis system. The technical details of the iMT-DDS are as follows:

1. *ROI definition and preprocessing*: We adopt the approach shown in Figure 3(a) for ROI definition to minimize the potential impact of border distortion. In addition, we adopt a single ROI min-max approach for the CEDM breast cancer study and a bi-ROI min-max approach for the CT cholangiocarcinoma study. Both case studies are presented later in this section.

2. *Texture feature extraction*: We adopt three texture algorithms—namely, GLCM, LBP, and Gabor filter. When applying these algorithms on a large image, the border pixels are just a small proportion of all pixels in the image; thus the distortion is not critical. For example, for a 100×100 image, 396 out of 10000 pixels (4%) are on the edges or corners

of the image so the impact of the border pixels may not be significant. However, when the size of image is 10×10 (like some ROIs in our study), 36 out of 100 pixels (36%) are border pixels, which requires some techniques to ameliorate this issue. Here, to minimize this border impact, we extend the ROI size to include real image pixels for texture feature calculation at the border (in contrast with the commonly used zero-padding approach that adds artificial all-dark pixels outside the window; this method would severely distort the calculation at the border). Then, when summarizing the texture feature calculation results into texture features, we only include the results on the pixels within the original image ROI. This approach turns out to produce superior performance over other alternative approaches in handling border distortion.

3. *Feature dimension reduction*: We use PCA to perform this function. To preserve clinical interpretability, we choose to apply PCA on each of the three texture algorithms individually. This accounts for the fact that the different texture algorithms describe the image from different perspectives, and so applying PCA on all the texture features together is inappropriate.

4. *Predictive modeling*: We adopt SVM, LDA, and QDA as the classifiers. In selecting features, the LOOCV is chosen as the objective function. SFS is applied for the breast cancer study, and PSO is used in the cholangiocarcinoma study.

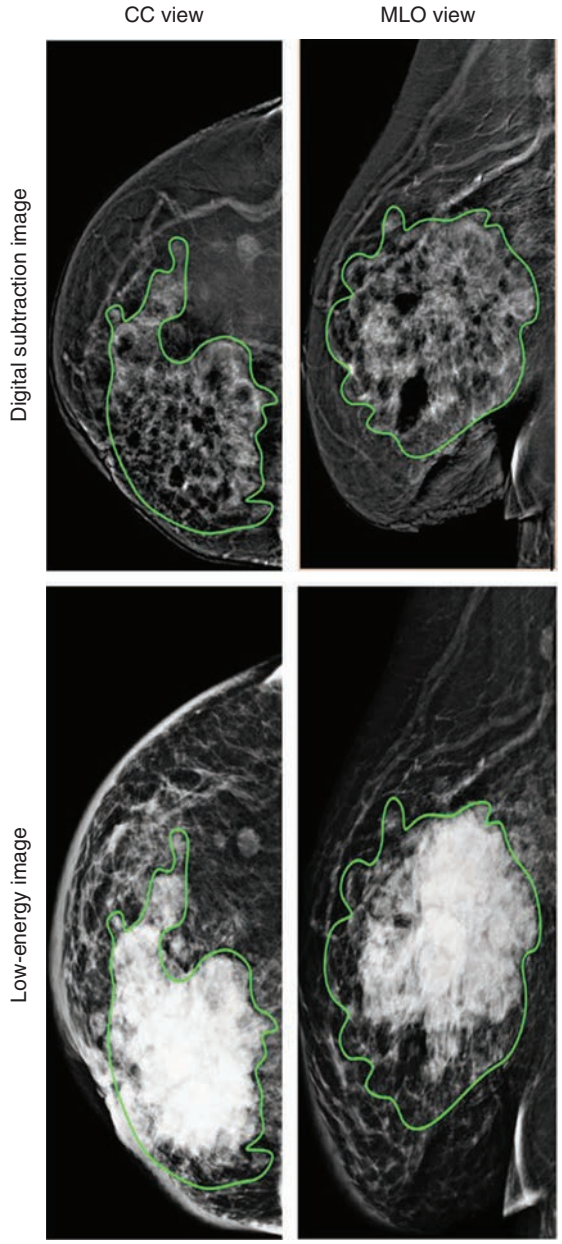
In the following sections, we present two case studies: breast cancer using CEDM imaging and cholangiocarcinoma using CT imaging. The first case study demonstrates the application of iMT-DDS to breast cancer diagnosis. The second case study shows that iMT-DDS is able to identify the association between imaging biomarkers with genomic driver targeting cholangiocarcinoma so that personalized treatment planning can be achieved.

4.1. Case Study I: Breast Cancer Diagnosis

Breast cancer is the second-leading cause of cancer death in women (Siegel et al. [54]) and accounts for about $\tilde{2}9\%$ of female cancer cases detected in the United States. But breast cancer is also one of the most treatable malignancies if it can be detected early. Three imaging techniques are commonly used clinically to diagnose breast cancer: ultrasound (US), DM, and MRI. US suffers from a high number of false positives, leading to excessive biopsies. MRI has good diagnosis performance but is costly. DM is much less expensive compared with MR, and it has an overall sensitivity of 75%–85%. But in high-risk patients with *BRCA* (breast cancer susceptibility gene) mutations or dense breasts, sensitivity decreases to the range 30%–50% (Kriege et al. [30], Kuhl et al. [31], Leach et al. [36], Warner et al. [61]). The lower sensitivity of mammography in women with dense breasts is most likely due to a masking effect caused by the large amount of fibroglandular tissue (Boyd et al. [5]). As a result, improved imaging techniques using contrast enhancement to detect breast cancer have been developed and refined for these high-risk patients, including dedicated breast CT (DBCT), breast MRI, and, more recently, CEDM. Efforts have also been spent on exploring the utilities of imaging texture features in improving the diagnosis accuracy. Tan et al. [58] propose the use of image features from four-view mammography for cancer diagnosis and achieve an AUC of 0.79. do Nascimento et al. [13] study the discrete wavelet transform on digital mammograms combined with a polynomial classifier and achieve a high accuracy, with an AUC of 0.98. Muramastu et al. [43] extensively study the radial local ternary patterns from mammograms for breast cancer diagnosis, and the highest performance model achieves an AUC of 0.90.

In this tutorial, we demonstrate the applicability of our iMT-DDS for breast cancer diagnosis using a novel imaging modality, CEDM. Note CEDM is an emerging technology for breast cancer diagnosis, and Mayo Clinic is the first Food and Drug Administration non-beta testing site to use this technology in clinical practice. CEDM generates a low-energy mammography image along with a recombined contrast-enhanced image reflecting contrast accumulation within a lesion (see Figure 9). Breast regions with increased and/or

FIGURE 9. CEDM images on a malignant breast cancer case.



Notes. Each lesion was captured in two views (CC versus MLO) with two images per view. One is a low-energy image, and the other is a digital subtraction image. The green contours are the lesion contours marked by the radiologist.

leaky vasculature, two common characteristics of neoplasms, can be identified using intravenously administered, iodinated contrast material. This additional information improves lesion detection and characterization. As seen in Figure 9, there are two views (cranial-caudal (CC) versus mediolateral oblique (MLO)) for the low-energy and contrast-enhanced energy. As a result, each patient will have four images and thus four ROIs to be included in the study.

Table 1. Parameter setting for texture analysis algorithms for case study I.

Algorithm	Parameter	Description	Guideline for selection
GLCM	Direction	In which direction the co-occurrence should be considered	$[0°, 45°, 90°, 135°]$
	Distance	How far between two points the co-occurrence should be considered	1
LBP	Direction	Radius of the neighborhood	1, 3
	Number of surrounding points	Number of points that should be selected on the circle of neighborhood	8, 24
	Number of bins of histogram	The granularity of the texture spectrum	10, 26 (Ojala et al. [45])
Gabor	Frequency	The frequency of the sinusoidal wave	Depends on the properties of images (0.1, 0.3, and 0.5 are used in our case)
	Direction	The direction of the sinusoidal wave	$[0°, 45°, 90°, 135°]$
	Variance	The variance of the Gaussian kernel	Depends on the properties of images (1 and 2 are used)

4.1.1. Data. In an Institutional Review Board (IRB) exempt study, we retrospectively reviewed contrast-enhanced spectral mammography examinations performed between August 2014 and December 2015 (Hologic, Bedford, Massachusetts). Informed consent was obtained from all patients having a BIRADS (Breast Imaging Reporting and Data Systems) 4 or 5 classification in a preexisting IRB-approved study determining whether CEDM could lower the false-positive biopsy rates in mammography. The patient cohort included examinations that met the following criteria: (1) diagnostic mammogram that received a BIRADS rating of 4 or 5 and (2) studies that corresponded with available pathology results from surgical or image-guided biopsy. We limited the cohort to BIRADS 4 and 5 lesions because the analysis required the gold standard of lesion pathology. In total, 99 patients were identified that met the above inclusion criteria (21 benign and 78 malignant biopsy-proven lesions); since this is an imbalanced data set, SMOTE (Chawla et al. [8]), a commonly adopted method for unbalanced data, was applied first before the experiments.

4.1.2. Experiments. Based on our previous exploration on various medical applications (Hu et al. [25, 26], Ramkumar et al. [47]), we set the parameters of GLCM, LBP, and Gabor as seen in Table 1. We apply the texture algorithm to each of the four views, resulting in 244 (61×4) texture features (61 texture features $= 13$ GLCM features $+ 36$ LBP features $+ 12$ Gabor features) for each patient. In addition, we have four intensity-based features: mean, standard deviation, skewness, and kurtosis.

We have two hypotheses in this study: (1) the texture features improve the diagnosis accuracy when compared with intensity-based features alone, and (2) contrast-enhanced imaging (from digital subtraction image (DES)) shows advantages over low-energy imaging for breast cancer diagnosis. To test the hypotheses, we design three experiments using LDA, QDA, and SVM methods. In the first experiment, we develop the classifiers from four intensity features derived from low-energy images. In the second experiment, we develop the classifiers on the intensity features and the 61 texture features, again from the low-energy images. In the third experiment, the feature space is expanded to include both low-energy and DES images, using both intensity and texture features. It is interesting to note that when using only intensity features (see Table 2), the performance of SVM is comparable to both LDA and QDA. However, its performance significantly deteriorates with texture

Table 2. Experimental results for breast cancer case study.

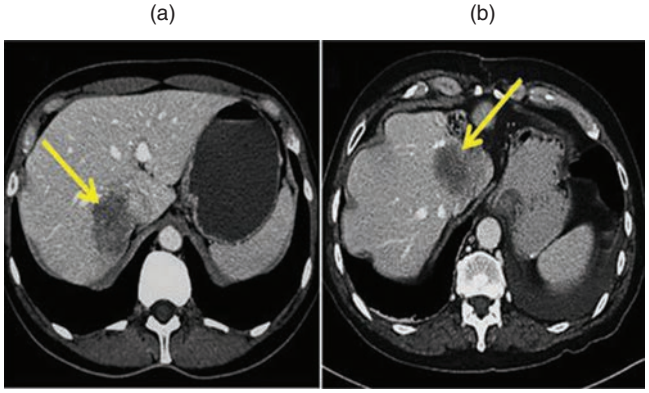
	Overall accuracy (%)	Sensitivity (%)	Specificity (%)
Experiment I			
LDA	74	73	76
QDA	77	73	90
SVM	77	74	86
Experiment II			
LDA	78	78	76
QDA	82	94	33
SVM	75	85	38
Experiment III			
LDA	82	82	81
QDA	87	91	71
SVM	78	96	10

features added into the model. Pending further investigation, we suspect the reason may be that the oversampling by SMOTE on the benign class does not help shift the classification boundary in SVM, such that the boundary is still to the benefit of the majority (i.e., malignant) class and makes the accuracy of minority class (i.e., specificity) low. This impact becomes more obvious when the feature dimension is high such as in experiments II and III when texture features are added. Nevertheless, LDA's performance supports the hypothesis well. Specifically, using the intensity-only feature gives an overall accuracy of 74% (sensitivity: 73%, specificity 76%); adding texture features from low energy improves the performance to accuracy 78% (sensitivity: 78%, specificity 76%). This demonstrates the value of texture features in cancer diagnosis. In addition, with added features (both intensity and texture) from DES images, the performance is further improved to an accuracy of 82% (sensitivity: 82%, specificity 81%).

4.2. Case Study II: Cholangiocarcinoma

Cholangiocarcinoma is an aggressive malignant tumor of the intrahepatic and extrahepatic bile ducts. It is very difficult to be detected, and as a result, patients usually have more advanced tumors by the time of diagnosis. The five-year survival rate only ranges from 2% to 30%. Currently, surgical excision and liver transplantation offer the only chance for curative therapy. Systemic therapy with gemcitabine hydrochloride and cisplatin is the current standard of care for patients with advanced disease. Unfortunately, this therapy provides only modest benefit, with dismal prospects for long-term survival (Valle et al. [59]). Driven by precision medicine, clinical researchers are looking into evaluation of particular genetic aberrations that are targetable by novel chemotherapeutic regimens. Fibroblast growth factor receptors (FGFRs) control biologic functions relevant to oncology, including cell survival, proliferation, migration, and differentiation. Several recent studies have implicated abnormal FGFR2 gene fusions in intrahepatic cholangiocarcinoma (Arai et al. [2], Borad et al. [4], Wu et al. [65]). While the identification of the FGFR2 has the potential of developing targeted therapeutic agents, the genomic analysis for diagnosis is invasive and costly. As a result, the use of imaging technology to noninvasively characterize tumors may have great potential to transform clinical practices. As shown in Figure 10, even an experienced radiologist has difficulty with differentiating FGFR2+ versus FGFR2− from the CT imaging. The central hypothesis of this study is that imaging texture features may be able to capture the subtle differences between FGFR2− and FGFR2+, and thus be able to differentiate between the two cohorts. If this hypothesis is shown to be true, we will be one step closer to precision medicine in this area.

FIGURE 10. CT imaging on liver: (a) FGFR2+ and (b) FGFR2−.



4.2.1. Data. Following institutional review board approval, a HIPAA–compliant retrospective study was performed. Thirty-three patients with advanced sporadic intrahepatic cholangiocarcinoma who had previously undergone integrated tumor sample whole-genome and whole-transcriptome analyses, as well as corresponding FGFR2 break-apart fluorescence in situ hybridization (FISH) assay, were evaluated with CT. Fifteen patients were positive for FGFR2 gene fusion, and the remaining 18 patients are FGFR2 negative.

4.2.2. Experiment. We set the parameters of GLCM, LBP, and Gabor as follows by referencing the parameter settings of similar applications (see Table 3). A total of 37 texture features (13, 12, and 12 from GLCM, LBP, and Gabor, respectively) are computed for each ROI. Thirty-three subjects are included in this study, with 15 FGFR2+ and 18 FGFR2− cases. The LOOCV classification accuracy is 91% (sensitivity = 87%, specificity = 94%; see Table 4).

One interesting observation from this study is all three methods have comparable performance similar to that from experiment I in the first case study, and SVM outperforms both

TABLE 3. Parameter setting for texture analysis algorithms for case study II.

Algorithm	Parameter	Description	Guideline for selection
GLCM	Direction	In which direction the co-occurrence should be considered	$[0°, 45°, 90°, 135°]$
	Distance	How far between two points the co-occurrence should be considered	1
LBP	Direction	Radius of the neighborhood	3
	Number of surrounding points	Number of points that should be selected on the circle of neighborhood.	24
	Number of bins of histogram	The granularity of the texture spectrum	12
Gabor	Frequency	The frequency of the sinusoidal wave	Depends on the properties of images (0.1, 0.3, and 0.5 are used in our case)
	Direction	The direction of the sinusoidal wave	$[0°, 45°, 90°, 135°]$
	Variance	The variance of the Gaussian kernel	Depends on the properties of images (1 and 2 are used)

Table 4. Experimental results for cholangiocarcinoma case study.

	Overall accuracy (%)	Sensitivity (%)	Specificity (%)
LDA	79	80	78
QDA	76	73	78
SVM	91	87	94

LDA and QDA in this case study. This may suggest that SVM is relatively more sensitive to the data set (balanced versus imbalanced) and the number of features when developing the model. In addition, these two case studies confirm that model performance varies depending on the data set, which is dependent on both imaging modality and disease type. As a result, it is recommended to have multiple classifiers included in the system such as our iMT-DDS for clinical applications. We also want to note that the second case study has a relatively small data set compared with the first one. PSO, an aggressive feature selection mechanism, is used for this small data set. Although the results are cross validated, the robustness of the model on a totally new data set for blind testing is yet to be explored.

5. Conclusions and Future Directions

Precision medicine is a nationwide initiative to uncover cancer pathways by taking into account individual differences in lifestyle, environment, and biology. It is expected to accelerate biomedical research and provide clinicians new tools to improve diagnosis accuracy and identify therapies that work best with individual patients. Recent studies have demonstrated the success of using imaging features in disease diagnosis (introducing the concept of radiomics) and identifying the associations between imaging features and genetics in various cancers (introducing the concept of radiogenomics). Aligning with these new concepts, quantitative imaging is becoming an emerging research field. The cornerstone of quantitative imaging research is a system that processes imaging for feature extraction in conjunction with multiparametric model development. In this tutorial, we provide a detailed discussion of the essential components of a quantitative imaging system. To illustrate the applicability of the system to clinical application, we present iMT-DDS, our in-house-developed imaging multitexture disease diagnosis system, on two clinical case studies. Since quantitative imaging is a relatively new field, here we offer some insight of future directions on each step of the overall process.

- Region of interest and normalization (step 1 of a quantitative imaging system): Manual segmentation of ROI has been the traditional method, but it is time consuming, tedious, and error-prone. In automated segmentation, some research focuses on intensity thresholding based on the assumption that tissues from ROIs have significantly and consistently different intensities than the background. These differences are either global, which only requires a fixed threshold, or local, which requires adaptive thresholding. In practice, however, the fundamental assumption of thresholding is often violated, and thresholding alone produces poor segmentation results. If at all, most segmentation methods apply thresholding only as a first step in the pipeline. Effort has also been devoted to developing a linear filter based on intensity-derived features rather than by their absolute intensities. Similar to thresholding, such feature-based filters alone usually do not produce definitive cell outlines but may provide useful cues for subsequent steps in the pipeline. Another popular class of filters includes those from the field of mathematical morphology. Being nonlinear, operators such as erosion, dilation, opening, and closing allow for the examination and manipulation of geometric and topological properties of objects in images, and they are often used in connection with cell segmentation. One common approach is to start from selected seed points in the image and to iteratively add connected points to form labeled regions. However, the morphological-based filtering is known to lack robust performance for images with large

noise and is infamous for producing oversegmentation. The fourth area is deformable model fitting where the models are formulated either explicitly, as the parametric contour (2D) or surface (3D), or implicitly, as the zero level of a function with dimensionality (nD) one higher than the image to be segmented. One popular method is level set; however, it requires computation of partial differential equations at each point on the iteratively evolving surface and thus has been known to be computationally prohibitive for widespread clinical use. Apparently, the challenging task in this step is the development of accurate, fast, and automated segmentation so that the ROI can be precisely delineated. Next, as stated earlier, normalization plays a significant role in the quality of the texture features to be extracted. A thorough investigation on disease and imaging modality-dependent normalization should be conducted so that a general principle on this procedure may be derived.

• Texture feature extraction (step 2 of a quantitative imaging system): Texture analysis is significantly impacted by the imaging resolution and quality (e.g., signal-to-noise ratio (SNR)). In MR imaging, the image resolution is defined by the field of view, slice thickness, and matrix, while the SNR is defined as the ratio of signal power to noise power. The higher the ratio, the stronger the signal compared to the noise. Apparently, texture feature discriminative power increases for higher-resolution images with higher SNR values. However, existing literature indicates that the texture application in clinical practice is far more complicated as good reproducibility across multicenter studies has not yet been achieved. More experimental studies on the quality of the texture features from different imaging are necessary. Second, most research (including ours) focuses on extracting texture features from 2D ROI. Extracting 3D texture features may further improve the clinical utility.

• Feature dimension reduction (step 3 of a quantitative imaging system) and predictive modeling (step 4 of a quantitative imaging system): Clinical data often have missing data. Often the portion of the data with missing values is just discarded, resulting in valuable clinical data not being fully utilized. Exploring ways of missing data imputation is becoming a popular topic in machine learning, and it deserves special attention for clinical applications. To avoid overfitting, research has separated the data into training and validation sets so that results on new data can be reported. This should be strictly applied to feature dimension reduction, feature selection, and predictive model development so that the results with robust predictive performance can be reported. However, most clinical studies are large p and small n problems—that is, the number of features (p) is much greater than the number of samples (patients, n), leading to challenges in generalized model development. How to take advantage of the larger existing publicly available data sets and develop/tune the predictive models to have robust performance is the next immediate question from this field.

Acknowledgments

The research presented in this tutorial was supported by Mayo Clinic Kern Center for Science of Healthcare Delivery and ASU–Mayo Clinic seed program.

Appendix. GLCM Features

Notation

To introduce the thirteen GLCM features, we need the following six basic notations.

$p(i,j)$: (i,j)th entry in a normalized gray-tone spatial-dependence matrix, $= P(i,j)/R$.
$p_x(i)$: ith entry in the marginal-probability matrix obtained by summing the rows of $p(i,j)$, $= \sum_{j=1}^{N_g} P(i,j)$.
N_g: Number of distinct gray levels in the quantized image.

$$p_y(j) = \sum_{i=1}^{N_g} P(i,j).$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P(i,j), \quad i+j=k, \textit{ for } k=2,3,\ldots,2N_g.$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P(i,j), \quad |i-j|=k, \textit{ for } k=0,1,\ldots,N_g-1.$$

The thirteen GLCM features are listed as follows.

(1) Angular second moment:

$$f_1 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\{P(i,j)\}^2.$$

(2) Contrast:

$$f_2 = \sum_{n=0}^{N_g-1} n^2\left\{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\{P(i,j)\}\right\}, \quad |i-j|=n.$$

(3) Correlation:

$$f_3 = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(ij)P(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y},$$

where μ_x, μ_y, σ_x, and σ_y are the means and standard deviations of ρ_x and ρ_y.

(4) Sum of squares: Variance:

$$f_4 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu)P(i,j).$$

(5) Inverse difference moment:

$$f_5 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\frac{1}{1+(i-j)^2}P(i,j).$$

(6) Sum average:

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i).$$

(7) Sum variance:

$$f_7 = \sum_{i=2}^{2N_g}(i-f_s)^2 p_{x+y}(i).$$

(8) Sum entropy:

$$f_8 = -\sum_{i=2}^{2N_g} p_{x+y}(i)\log\{p_{x+y}(i)\}.$$

(9) Entropy:

$$f_9 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log\{p(i,j)\}.$$

(10) Difference variance:

$$f_{10} = \text{variance of } p_{x-y}.$$

(11) Difference entropy:

$$f_{11} = -\sum_{i=0}^{N_g-1} p_{x-y}(i)\log\{p_{x-y}(i)\}.$$

(12), (13) Information measures of correlation:

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}.$$

$$f_{13} = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2}.$$

$$HXY = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log\{p(i,j)\}.$$

$$HXY1 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log\{p_x(i)p_y(j)\}.$$

$$HXY2 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p_x(i)p_y(j)\log\{p_x(i)p_y(j)\}.$$

References

[1] H. K. Angell, N. Gray, C. Womack, D. I. Pritchard, R. W. Wilkinson, and M. Cumberbatch. Digital pattern recognition-based image analysis quantifies immune infiltrates in distinct tissue regions of colorectal cancer and identifies a metastatic phenotype. *British Journal of Cancer* 109(6):1618–1624, 2013.

[2] Y. Arai, Y. Totoki, F. Hosoda, T. Shirota, N. Hama, H. Nakamura, H. Ojima, et al. Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma. *Hepatology* 59(4):1427–1434, 2014.

[3] C. Atupelage, H. Nagahashi, M. Yamaguchi, T. Abe, A. Hashiguchi, and M. Sakamoto. Computational grading of hepatocellular carcinoma using multifractal feature description. *Computerized Medical Imaging and Graphics* 37(1):61–71, 2013.

[4] M. J. Borad, M. D. Champion, J. B. Egan, W. S. Liang, R. Fonseca, A. H. Bryce, A. E. McCullough, et al. Integrated genomic characterization reveals novel, therapeutically relevant drug targets in FGFR and EGFR pathways in sporadic intrahepatic cholangiocarcinoma. *PLoS Genetics* 10(2):e1004135, 2014.

[5] N. F. Boyd, H. Guo, L. J. Martin, L. Sun, J. Stone, E. Fishell, R. A. Jong, et al. Mammographic density and the risk and detection of breast cancer. *New England Journal of Medicine* 356(3): 227–236, 2007.

[6] L. Cai, X. Wang, Y. Wang, Y. Guo, J. Yu, and Y. Wang. Robust phase-based texture descriptor for classification of breast ultrasound images. *Biomedical Engineering Online* 14: Article 26, 2015.

[7] A. Chaddad, C. Desrosiers, A. Bouridane, M. Toews, L. Hassan, and C. Tanougast. Multi texture analysis of colorectal cancer continuum using multispectral imagery. *PLoS One* 11(2):e0149893, 2016.

[8] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* 16(1):321–357, 2002.

[9] X. Chen, X. Wei, Z. Zhang, R. Yang, Y. Zhu, and X. Jiang. Differentiation of true-progression from pseudoprogression in glioblastoma treated with radiation therapy and concomitant temozolomide by GLCM texture analysis of conventional MRI. *Clinical Imaging* 39(5):775–780, 2015.

[10] C. D. Chong, N. Gaw, Y. Fu, J. Li, T. Wu, and T. J. Schwedt. Migraine classification using magnetic resonance imaging resting-state functional connectivity data. *Cephalalgia* 37(9): 828–844, 2017

[11] G. Collewet, M. Strzelecki, and F. Mariette. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magnetic Resonance Imaging* 22(1):81–91, 2004.

[12] F. Davnall, C. S. Yip, G. Ljungqvist, M. Selmi, F. Ng, B. Sanghera, B. Ganeshan, K. A. Miles, G. J. Cook, and V. Goh. Assessment of tumor heterogeneity: An emerging imaging tool for clinical practice. *Insights into Imaging* 3(6):573–589, 2012.

[13] M. Z. do Nascimento, A. S. Martins, L. A. Neves, R. P. Ramos, E. L. Flores, and G. A. Carrijo. Classification of masses in mammographic image using wavelet domain features and polynomial classifier. *Expert Systems with Applications* 40(15):6213–6221, 2013.

[14] G. Doumou, M. Siddique, C. Tsoumpas, V. Goh, and G. J. Cook. The precision of textural analysis in (18)F-FDG-PET scans of oesophageal cancer. *European Radiology* 25(9):2805–2812, 2015.

[15] J. M. Dunn, T. Hveem, M. Pretorius, D. Oukrif, B. Nielsen, F. Albregtsen, L. B. Lovat, M. R. Novelli, and H. E. Danielsen. Comparison of nuclear texture analysis and image cytometric DNA analysis for the assessment of dysplasia in Barretts oesophagus. *British Journal of Cancer* 105(8):1218–1223, 2011.

[16] Y. Fan, S. M. Resnick, X. Wu, and C. Davatzikos. Structural and functional biomarkers of prodromal Alzheimers disease: A high-dimensional pattern classification study. *Neuroimage* 41(2):277–85, 2008.

[17] M. Field, T. F. Witham, J. C. Flickinger, D. Kondzioka, and L. D. Lunsford. Comprehensive assessment of hemorrhage risks and outcomes after stereotactic brain biopsy. *Journal of Neurosurgery* 94(4):545–551, 2001.

[18] M. J. Gangeh, H. Tadayyon, L. Sannachi, A. Sadeghi-Naini, W. T. Tran, and G. J. Czarnota. Computer aided theragnosis using quantitative ultrasound spectroscopy and maximum mean discrepancy in locally advanced breast cancer. *IEEE Transactions on Medical Imaging* 35(3):778–790, 2016.

[19] N. Gaw, T. J. Schwedt, C. D. Chong, T. Wu, and J. Li. A clinical decision support system using multi-modality imaging data for disease diagnosis. *IIE Transactions on Healthcare Systems Engineering*, 2017. Forthcoming.

[20] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research* 3(March):1157–1182, 2003.

[21] F. Han, H. Wang, G. Zhang, H. Han, B. Song, L. Li, W. Moore, H. Lu, H. Zhao, and Z. Liang. Texture feature analysis for computer-aided diagnosis on pulmonary nodules. *Journal of Digital Imaging* 28(1):99–115, 2015.

[22] A. N. Hanannia, L. E. Bantis, Z. Feng, H. Wang, E. P. Tamm, M. H. Katz, A. Maitra, and E. J. Koay. Quantitative imaging to evaluate malignant potential of IPMNs. *Oncotarge* 7(52): 85776–85784, 2016.

[23] R. M. Haralick and K. Shanmugam. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* 3(6):610–621, 1973.

[24] L. C. Harrison, R. Nikander, M. Siki, T. Luukkaala, M. T. Helminen, P. Ryymin, S. Soimakallio, H. J. Eskola, P. Dastidar, and H. Sievnen. MRI texture analysis of femoral neck: Detection of exercise load-associated differences in trabecular bone. *Journal of Magnetic Resonance Imaging* 34(6):1359–1366, 2011.

[25] L. S. Hu, S. Ning, J. M. Eschbacher, L. C. Baxter, N. Gaw, S. Ranjbar, J. Plasencia, et al. Radiogenomics to characterize regional genetic heterogeneity in glioblastoma. *Neuro-Oncology* 19(1):128–137, 2017.

[26] L. S. Hu, S. Ning, J. M. Eschbacher, N. Gaw, A. C. Dueck, K. A. Smith, P. Nakaji, et al. Multi-parametric MRI and texture analysis to visualize spatial histologic heterogeneity and tumor extent in glioblastoma. *PLoS One* 10(11):e0141506, 2015.

[27] S. Huang, J. Li, J. Ye, T. Wu, K. Chen, A. Fleisher, and E. Reiman. Identifying Alzheimers disease-related brain regions from multi-modality neuroimaging data using sparse composite linear discrimination analysis. J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett F. Pereira, and K. Q. Weinberger, eds. *Advances in Neural Information Processing Systems 24 (NIPS 2011)*. Curran Associates, Red Hook, NY, 1431–1439, 2011.

[28] R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis*, 5th ed. Prentice Hall, Upper Saddle River, NJ, 2002.

[29] A. Kassner and R. E. Thornhill. Texture analysis: A review of neurologic MR imaging applications. *American Journal of Neuroradiology* 31(5):809–816, 2010.

[30] M. Kriege, C. T. Brekelmans, C. Boetes, P. E. Besnard, H. M. Zonderland, I. M. Obdeijn, R. A. Manoliu, et al. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. *New England Journal of Medicine* 351(5):427–437, 2004.

[31] C. K. Kuhl, R. K Schmutzler, C. C Leutner, A. Kempe, E. Wardelmann, A. Hocke, M. Maringa, U. Pfeifer, D. Krebs, and H. H. Schild. Breast MR imaging screening in 192 women proved or suspected to be carriers of a breast cancer susceptibility gene: Preliminary results 1. *Radiology* 215(1):267–279, 2000.

[32] A. Larroza, V. Bodi, and D. Moratal. Texture analysis in magnetic resonance imaging: Review and considerations for future applications. C. Constantinides, ed. *Assessment of Cellular and Organ Function and Dysfunction Using Direct and Derived MRI Methodologies*, Chapter 5. InTech, Rijeka, Croatia, 2016.

[33] A. Lashkari. Early breast cancer detection in thermogram images using supervised and unsupervised algorithms. *Middle East Journal of Cancer* 7(3):113–124, 2016.

[34] A. Laurinavicius, A. Laurinaviciene, D. Dasevicius, N. Elie, B. Plancoulaine, C. Bor, and P. Herlin. Digital image analysis in pathology: Benefits and obligation. *Analytical Cellular Pathology* 35(2):75–78, 2012.

[35] A. Laurinavicius, B. Plancoulaine, A. Laurinaviciene, P. Herlin, R. Meskauskas, I. Baltrusaityte, J. Besusparis, et al. A methodology to ensure and improve accuracy of Ki67 labelling index estimation by automated digital image analysis in breast cancer tissue. *Breast Cancer Research* 16(2):R35, 2014.

[36] M. O. Leach, C. R. Boggis, A. K. Dixon, D. F. Easton, R. A. Eeles, D. G. Evans, F. J. Gilbert, et al. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: A prospective multicentre cohort study (MARIBS). *Lancet* 365(9473):1769–1778, 2005.

[37] P. Lo, S. Young, H. J. Kim, M. S. Brown, and M. F. McNitt-Gray. Variability in CT lung-nodule quantification: Effects of dose reduction and reconstruction methods on density and texture based features. *Medical Physics* 43(8):4854–4865, 2016.

[38] C. P. Loizou, S. Petroudi, I. Seimenis, M. Pantziaris, and C. S Pattichis. Quantitative texture analysis of brain white matter lesions derived from T2-weighted MR images in MS patients with clinically isolated syndrome. *Journal of Neuroradiology* 42(2):99–114, 2015.

[39] W. Lu, Z. Li, and J. Chu. A novel computer-aided diagnosis system for breast MRI based on feature selection and ensemble learning. *Computers in Biology and Medicine* 83:157–165, 2017.

[40] A. Materka. Texture analysis methodologies for magnetic resonance imaging. *Dialogues in Clinical Neuroscience* 6(2):243–250, 2004.

[41] A. Materka, M. Strzelecki, R. Lerski, and L. Schad. Feature evaluation of texture test objects for magnetic resonance imaging. *Series in Machine Perception and Artificial Intelligence* 40: 197–206, 2000.

[42] N. Michoux, S. Van den Broeck, L. Lacoste, L. Fellah, C. Galant, M. Berlire, and I. Leconte. Texture analysis on MR images helps predicting non-response to NAC in breast cancer. *BMC Cancer* 15(1):Article 574, 2015.

[43] C. Muramatsu, T. Hara, T. Endo, and H. Fujita. Breast mass classification on mammograms using radial local ternary patterns. *Computers in Biology and Medicine* 72:43–53, 2016.

[44] National Cancer Institute. Understanding cancer: Cancer statistics. Accessed March 22, https://www.cancer.gov/about-cancer/understanding/statistics, 2017.

[45] T. Ojala, M. Pietikainen, and T. Maenpaa. A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. S. Singh, N. A. Murshed, and W. G. Kropatsch, eds. *Proceedings of the Second International Conference on Advances in Pattern Recognition* (*ICAPR '01*), Springer, Berlin, 397–406, 2001.

[46] T. D. Pham, Y. Watanabe, M. Higuchi, and H. Suzuki. Texture analysis and synthesis of malignant and benign mediastinal lymph nodes in patients with lung cancer on computed tomography. *Scientific Reports* 7:43209, 2017.

[47] S. Ramkumar, S. Ranjbar, S. Ning, D. Lal, C. M. Zwart, C. P. Wood, S. M. Weindling, et al. MRI-based texture analysis to differentiate sinonasal squamous cell carcinoma from inverted papilloma. *American Journal of Neuroradiology* 38(5):1019-1025, 2017.

[48] N. K. Ratha, A. K. Jain, and S. Lakshmanan. Object detection using Gabor filters. *Pattern Recognition* 30(2):295–309, 1997.

[49] Y. A. Reyad, M. A. Berbar, and M. Hussain. Comparison of statistical, LBP, and multi-resolution analysis features for breast mass classification. *Journal of Medical Systems* 38(9): Article 100, 2014.

[50] Candy Sagon. Cancer still top killer, but death rate is dropping. Accessed August 22, 2017, http://blog.aarp.org/2014/01/09/cancer-still-top-killer-but-death-rate-is-dropping/?intcmp=AE-BLIL-BL, 2017.

[51] M. Sanz-Cortés, F. Figueras, E. Bonet-Carne, N. Padilla, V. Tenorio, N. Bargalló, I. Amat-Roldan, and E. Gratacs. Fetal brain MRI texture analysis identifies different microstructural patterns in adequate and small for gestational age fetuses at term. *Fetal Diagnosis and Therapy* 33(2):122–129, 2013.

[52] T. J. Schwedt, C. D. Chong, T. Wu, N. Gaw, Y. Fu, and J. Li. Accurate classification of chronic migraine via brain magnetic resonance imaging. *Headache: Journal of Head and Face Pain* 55(6):762–777, 2015.

[53] Z. Shi, Z. Yang, G. Zhang, G. Cui, X. Xiong, Z. Liang, and H. Lu. Characterization of texture features of bladder carcinoma and the bladder wall on MRI: Initial experience. *Academic Radiology* 20(8):930–938, 2013.

[54] R. L. Siegel, K. D. Miller, and A. Jemal. Cancer statistics, 2016. *CA: Cancer Journal for Clinicians* 66(1):7–30, 2016.

[55] K. Skogen, B. Ganeshan, C. Good, G. Critchley, and K. Miles. Measurements of heterogeneity in gliomas on computed tomography relationship to tumour grade. *Journal of Neuro-Oncology* 111(2):213–219, 2013.

[56] G. Song, F. Xue, and C. Zhang. A model using texture features to differentiate the nature of thyroid nodules on sonography. *Journal of Ultrasound in Medicine* 34(10):1753–1760, 2015.

[57] S. S. Suganthi and S. Ramakrishnan. Analysis of breast thermograms using Gabor wavelet anisotropy index. *Journal of Medical Systems* 38(9):Article 101, 2014.

[58] M. Tan, J. Pu, and B. Zheng. Reduction of false-positive recalls using a computerized mammographic image feature analysis scheme. *Physics in Medicine and Biology* 59(15):4357–4573, 2014.

[59] J. Valle, H. Wasan, D. H. Palmer, D. Cunningham, A. Anthoney, A. Maraveyas, S. Madhusudan, et al. Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer. *New England Journal of Medicine* 362(14):1273–1281, 2010.

[60] P. Wang, W. Xu, J. Sun, C. Yang, G. Wang, Y. Sa, X. H. Hu, and Y. Feng. A new assessment model for tumor heterogeneity analysis with [18]F-FDG PET images. *EXCLI Journal* 15:75–84, 2016.

[61] E. Warner, D. B. Plewes, K. A. Hill, P. A. Causer, J. T. Zubovits, R. A. Jong, M. R. Cutrara, et al. Surveillance of *BRCA1* and *BRCA2* mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. *Journal of the American Medical Association* 292(11):1317–1325, 2004.

[62] A. W. Whitney. A direct method of nonparametric measurement selection. *IEEE Transactions on Computers* 100(9):1100–1103, 1971.

[63] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation* 1(1):67–82, 1997.

[64] W. Wu, C. Parmar, P. Grossmann, J. Quackenbush, P. Lambin, J. Bussink, R. Mark, and H. J. W. Aerts. Exploratory study to identify radiomics classifiers for lung cancer histology. *Frontiers in Oncology* 6: Article 71, 2016.

[65] Y. M. Wu, F. Su, S. Kalyana-Sundaram, N. Khazanov, B. Ateeq, X. Cao, R. J. Lonigro, et al. Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discovery* 3(6): 636–647, 2013.

[66] H. Yang, J. Liu, J. Sui, G. Pearlson, and V. D. Calhoun VD. A hybrid machine learning method for fusing fMRI and genetic data: Combining both improves classification of schizophrenia. *Frontiers in Human Neuroscience* 4:192, 2010.

[67] S. H. Yoon, Y. H. Kim, Y. J. Lee, J. Park, J. W. Kim, H. S. Lee, and B. Kim. Tumor heterogeneity in human epidermal growth factor receptor 2 (HER2)-positive advanced gastric cancer assessed by CT texture analysis: Association with survival after trastuzumab treatment. *PLoS One* 11(8):e0161278, 2016.

[68] E. I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E. R. Melhem, and C. Davatzikos. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magnetic Resonance in Medicine* 62(6):1609–1618, 2009.

[69] D. Zhang, Y. Wang, L. Zhou, H. Yuan, D. Shen, and Alzheimer's Disease Neuroimaging Initiative. Multimodal classification of Alzheimers disease and mild cognitive impairment. *Neuroimage* 55(3):856–867, 2011.

[70] Q. Zhang, Q. Wu, J. Zhang, L. He, J. Huang, J. Zhang, H. Huang, and Q. Gong. Discriminative analysis of migraine without aura: Using functional and structural MRI with a multi-feature classification approach. *PLoS One* 11(9):e0163875, 2016.

[71] X. Zhang, X. Xu, Q. Tian, B. Li, Y. Wu, Z. Yang, Z. Liang, Y. Liu, G. Cui, and H. Lu. Radiomics assessment of bladder cancer grade using texture features from diffusion-weighted imaging. *Journal of Magnetic Resonance Imaging*, ePub ahead of print February 15, http://dx.doi.org/10.1002/jmri.25669, 2017.

[72] P. O. Zinn, B. Majadan, P. Sathyan, S. Majumder S. K. Singh, F. A. Jolesz, and R. R. Colen. Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. *PLoS One* 6(10):e25451, 2011.