A hierarchical spectral clustering and nonlinear dimensionality reduction scheme for detection of prostate cancer from magnetic resonance spectroscopy (MRS)

Pallavi Tiwari
Department of Biomedical Engineering, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854

Mark Rosen
Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Anant Madabhushia
Department of Biomedical Engineering, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854

(Received 10 November 2008; revised 26 June 2009; accepted for publication 26 June 2009; published 12 August 2009)

Magnetic resonance spectroscopy (MRS) has been shown to have great clinical potential as a supplement to magnetic resonance imaging in the detection of prostate cancer (CaP). MRS provides functional information in the form of changes in the relative concentration of specific metabolites including choline, creatine, and citrate which can be used to identify potential areas of CaP. With a view to assisting radiologists in interpretation and analysis of MRS data, some researchers have begun to develop computer-aided detection (CAD) schemes for CaP identification from spectroscopy. Most of these schemes have been centered on identifying and integrating the area under metabolite peaks which is then used to compute relative metabolite ratios. However, manual identification of metabolite peaks on the MR spectra, and especially via CAD, is a challenging problem due to low signal-to-noise ratio, baseline irregularity, peak overlap, and peak distortion. In this article the authors present a novel CAD scheme that integrates nonlinear dimensionality reduction (NLDR) with an unsupervised hierarchical clustering algorithm to automatically identify suspicious regions on the prostate using MRS and hence avoids the need to explicitly identify metabolite peaks. The methodology comprises two stages. In stage 1, a hierarchical spectral clustering algorithm is used to distinguish between extracapsular and prostatic spectra in order to localize the region of interest (ROI) corresponding to the prostate. Once the prostate ROI is localized, in stage 2, a NLDR scheme, in conjunction with a replicated clustering algorithm, is used to automatically discriminate between three classes of spectra (normal appearing, suspicious appearing, and indeterminate). The methodology was quantitatively and qualitatively evaluated on a total of 18 1.5 T in vivo prostate T2-weighted (w) and MRS studies obtained from the multisite, multi-institutional American College of Radiology (ACRIN) trial. In the absence of the precise ground truth for CaP extent on the MR imaging for most of the ACRIN studies, probabilistic quantitative metrics were defined based on partial knowledge on the quadrant location and size of the tumor. The scheme, when evaluated against this partial ground truth, was found to have a CaP detection sensitivity of 89.33% and specificity of 79.79%. The results obtained from randomized threefold and fivefold cross validation suggest that the NLDR based clustering scheme has a higher CaP detection accuracy compared to such commonly used MRS analysis schemes as z score and PCA. In addition, the scheme was found to be robust to changes in system parameters. For 6 of the 18 studies an expert radiologist laboriously labeled each of the individual spectra according to a five point scale, with 1/2 representing spectra that the expert considered normal and 3/4/5 being spectra the expert deemed suspicious. When evaluated on these expert annotated datasets, the CAD system yielded an average sensitivity (cluster corresponding to suspicious spectra being identified as the CaP class) and specificity of 81.39% and 64.71%, respectively. © 2009 American Association of Physicists in Medicine. [DOI: 10.1118/1.3180955]

Key words: magnetic resonance spectroscopy, computer-aided diagnosis, prostate cancer, nonlinear dimensionality reduction, hierarchical clustering, unsupervised classification, k-means, replicated clustering, principal component analysis, locally linear embedding, graph embedding, spectral clustering, z score
Prostatic adenocarcinoma (CaP) is the most commonly occurring malignancy among men with 186,320 new cases and 28,660 deaths estimated to occur in the United States in 2008 (American Cancer Society, 2008). Early detection of CaP offers the best hope of curing it; however, early prostate cancer is usually asymptomatic. Screening of CaP is based on digital rectal examination and monitoring elevated levels of the blood serum prostate specific antigen (PSA). Definitive diagnosis of CaP involves histological examination of biopsy specimens obtained via a blinded sextant transrectal ultrasound (TRUS) directed biopsy for patients with elevated PSA levels. Since prostate ultrasound is limited in its ability to identify CaP, biopsy locations are chosen at random within the prostate sextants. Consequently, the CaP detection accuracy associated with TRUS is only 20%–25% in patients with elevated PSA levels (4–10 μg/ml). Recently, in vivo endorectal T2-weighted (w) structural magnetic resonance (MR) imaging (MRI) of the prostate has allowed for greater discrimination between benign and cancerous prostatic structures as compared to TRUS. However, structural T2-w MRI by itself has been shown to be limited in its ability to detect small foci of carcinoma contributing to a relatively low detection specificity.

Over the past few years, MR spectroscopic (MRS) imaging (MRSI) has emerged as a useful complement to structural MRI for potential screening of CaP. MRSI is a non-invasive technique used to obtain the metabolic concentrations of specific molecular markers and biochemicals in the prostate including citrate, creatine, and choline, and changes in the ratio [choline/citrate or (choline+creatine)/citrate] of which have been shown to be linked to presence of CaP. The spectra are obtained at either single or multiple locations from a rectangular spectral grid placed on a corresponding T2-w MR image. It has been demonstrated previously that the relative concentrations of choline, citrate, and creatine are significantly different in CaP and normal regions within the prostate.

The relative concentrations of choline, creatine, and citrate are obtained by calculating the area under the peak for these metabolites to assess the presence of CaP at a specific prostate location on the T2-w MRI. Identification of precise location of specific metabolites on the MR spectra is a difficult task for radiologists due to (a) a low signal to noise ratio and (b) the presence of biomedical signal artifacts associated with MR spectra such as peak overlap and peak and baseline distortion. Figure 1(a) shows a MRS grid superimposed on the corresponding T2-w MRI slice, while Fig. 1(b) shows the MR spectra acquired from individual MRS voxels from the grid shown in Fig. 1(a). Figures 1(c)–1(e) correspond to representative CaP, noisy and normal spectra obtained from three different locations (shown in red, blue, and green, respectively) within the spectral grid [Fig. 1(b)]. Note the amount of noise in the MRS in Fig. 1(d) which, in some cases, can severely limit the ability of a radiologist to accurately identify individual metabolite peaks and quantify corresponding peak areas. Thus the usefulness of MRSI as an adjunct to MRI as a means of detecting, localizing, and characterizing CaP is highly dependent on the quality of the spectral examinations obtained, owing to the challenges in visually identifying spectroscopic CaP signatures (through the identification of abnormal metabolite peak area ratios) with poor data quality.

Computer-aided diagnosis (CAD) in medical imaging serves as an adjunctive method for image or data analysis, aiding the radiologist in identifying areas of disease or abnormality. The low detection accuracy associated with current TRUS prostate biopsies points to a need for developing an image guided decision support system to direct needle placement in the prostate to increase CaP detection accuracy. While our group and others have begun to develop CAD systems for CaP detection from structural and functional MR imaging, corresponding developments in MRS have not been as forthcoming. This is in spite of evidence to suggest that integration of structural and metabolic imaging could boost diagnostic yield over that which could be obtained from any individual modality. To date, computer-based approaches for prostate MRSI analysis have focused largely on the use of semiautomated peak area inte-
tion to determine metabolite ratios, although a few researchers\textsuperscript{17–21} explored the use of more sophisticated techniques for automated peak finding.

Previous attempts at computerized analysis of MRS can be classified into two broad categories: (a) signal quantification (model dependent) and (b) pattern recognition (model independent) approaches. Commonly used quantification methods include VARPRO,\textsuperscript{17} AMARES,\textsuperscript{18} and QUEST,\textsuperscript{19} which are software utilities where the objective is to minimize the squared distance between the acquired data and a model basis function built on prior knowledge about the metabolic profiles of a typical MR spectrum. However, the performance of these quantification models is usually dependent on (i) the choice of correct number of model components, (ii) optimal choice of prior knowledge (model function), (iii) presence of noise and contributions from nonprostate spectra, (iv) peak overlap owing to contributions from multiple metabolites, and (v) baseline distortion and line broadening.\textsuperscript{22} In order to avoid the limitations of model and peak detection based approaches for MRS, recently some researchers have begun to explore domain independent techniques such as z score and principal component analysis (PCA). An excellent comprehensive comparison of quantification and pattern recognition schemes used for MRS analysis is provided in Ref.\textsuperscript{22}.

z score is a statistical quantity obtained as the ratio of the difference in each individual sample’s data and the population mean to the population variance, z score analysis\textsuperscript{20,23} aims to quantify the total contribution of all metabolites in the spectral vector. In Refs.\textsuperscript{20} and \textsuperscript{23}, z score was employed for MRS based detection of glioma and CaP, respectively. Another statistical technique, canonical correlation analysis (CCA),\textsuperscript{24} based on calculating the canonical coefficients to obtain correlated linear relationships between two multidimensional variables was shown to be useful in successfully classifying prostate MRSI datasets into four classes: Aggressive tumor, tumor, mixed tissue, and healthy tissue.

Devos et al.\textsuperscript{25} used linear discriminant analysis (LDA) and a least squares support vector machine (SVM) classifier to discriminate between different tumor classes on brain MRS. Ma and Sun\textsuperscript{26} and Simonetti et al.\textsuperscript{27} explored other linear dimensionality reduction methods such as independent component analysis and PCA, in conjunction with support vector machine (SVM) classifier, to differentiate between brain tissue classes via MRS.

Dimensionality reduction (DR) refers to the projection of high dimensional data into a reduced dimensional feature space without a significant loss in class discriminatory information. The low dimensional representation of the data is easier to visualize and DR algorithms aim to preserve object relationships so that objects that are close to one another in the high dimensional ambient space are mapped to adjacent locations in the resulting low dimensional embedding space. However, linear DR schemes such as PCA assume the original high dimensional data to be inherently linear and hence employ linear projection methods to reduce data dimensionality. Recently several nonlinear dimensionality reduction (NLDR) algorithms have been proposed for the analysis and visualization of nonlinear data.\textsuperscript{29–31} The objective behind NLDR methods is to nonlinearly map objects, c, d belonging to the same object class and characterized by M-dimensional feature vectors $F(c), F(d)$, to adjacent locations $S(c), S(d)$ in the low dimensional embedding: $S(c), S(d)$ representing the $m$-dimensional dominant eigenvectors corresponding to $c, d (m < M)$. Unlike linear DR schemes that typically employ the Euclidean distance measure to estimate object distances, most NLDR schemes aim at preserving geodesic distances between objects while computing the embedding. We have previously demonstrated that the use of NLDR methods for representation of high dimensional gene- and protein-expression data, results in better classification compared to linear DR based representation (PCA and LDA).\textsuperscript{28}

In this paper, we present a fully automated CAD system for detecting abnormal/suspicious regions on the prostate using 1.5 T prostate MRSI data. Figure 2 illustrates the organization of our CAD scheme. Our methodology comprises of two stages. In stage 1, a novel hierarchical classification scheme is employed to recursively distinguish prostatic from extracapsular (noninformative) spectra via graph embedding (GE),\textsuperscript{31,32} a well known NLDR scheme, in order to hone in the region of interest (ROI) corresponding to the prostate. In a typical field of view (FOV) of an in vivo endorectal prostate T2-w MR image, the prostate occupies a small percentage [approximately 10% in Fig. 1(a)] of the total volume within a prostate MRI scene. Since the extracapsular spectra are most populous, the largest cluster is identified at each iteration as being noninformative and eliminated. This process is repeated until the number of MR spectra remaining is approximately equal to the number usually contained in the prostate ($\Theta$) [Fig. 1(a)]. The removal of extracapsular spectra in stage 1 makes it easier to discriminate between suspicious appearing and normal appearing spectra within the prostate.

In stage 2, a NLDR scheme is applied to nonlinearly embed the informative spectra into a reduced dimensional space. The individual prostate spectra now characterized by their low dimensional embedding coordinates are clustered into distinct classes via a “replicated clustering” scheme. All spectra are aggregated into three classes based on the assumption that they correspond to normal, CaP, and indeterminate classes.

![Flowchart showing various system components and methodological overview of the CAD scheme. Hierarchical spectral clustering is performed in stage 1 to automatically obtain the prostate grid, followed in stage 2 by the identification of different tissue classes via NLDR and replicated clustering.](image-url)
The rest of this paper is organized as follows: In Sec. II we provide a detailed description of the feature extraction and clustering schemes employed in this paper. In Sec. III we provide the details of our CAD methodology for prostate MRS classification. In Sec. IV we explain the evaluation scheme employed in this work followed in Sec. V by results of qualitative and quantitative evaluation of our scheme on a total of 18 prostate MRI-MRS studies. Concluding remarks and future research directions are presented in Sec. VI.

II. DESCRIPTION OF FEATURE EXTRACTION AND REPLICATED CLUSTERING METHODS

II.A. Feature extraction methods

II.A.1. Linear dimensionality reduction scheme: PCA

PCA is a linear DR method widely used to visualize high dimensional data and discern object relationships in the data by finding orthogonal axes that contain the greatest amount of variance in the data. These orthogonal eigenvectors corresponding to the largest eigenvalues are called “principal components.” To obtain these principal components, each data point \( c \) in set \( C \) is first centered by subtracting the mean of all the features for each observation \( c \) from its original \( M \)-dimensional feature value \( f_{i}(c), u \in \{1, \ldots, M\} \) as shown by

\[
\tilde{f}_{u}(c) = f_{u}(c) - \left( \frac{1}{|C|} \sum_{c \in C} f_{u}(c) \right), \quad u \in \{1, \ldots, M\}. \tag{2.1}
\]

From feature values \( \tilde{f}(c) \) for each \( c \in C \), a new \( |C| \times M \) matrix \( \mathcal{Y} \) is constructed, where \(|C|\) is the cardinality of set \( C \). The matrix \( \mathcal{Y} \) is then decomposed into corresponding singular values as shown by

\[
\mathcal{Y} = U W_{\text{PCA}} V^{T}, \tag{2.2}
\]

where via singular value decomposition a \( |C| \times |C| \) diagonal matrix \( W_{\text{PCA}} \) containing the eigenvalues of the principal components, a \( M \times |C| \) left singular matrix \( U \), and a \( |C| \times M \) matrix \( V \) are obtained. The eigenvalues in \( W_{\text{PCA}} \) represent the amount of variance for each eigenvector \( S_{\text{PCA}} \), \( \nu \in \{1,2,\ldots,m\} \), in matrix \( V \) and are used to rank the corresponding eigenvectors in the order of greatest variance. Thus the first \( m \) eigenvectors that represent a prespecified percentage of the variance in the data are extracted, while the remaining Eigen vectors are discarded. Thus each data sample \( c \in C \) is now described by an \( M \)-dimensional embedding vector \( S_{\text{PCA}}(c) \). In spite of the fact that PCA assumes that the data lie on a linear manifold, it allows for specification of the number of eigenvectors required to explain a prespecified percentage of the variance in the data.

II.A.2. z score

For a set of objects, \( C \subset C \subset C \), a mean vector \( F_{\mu} = \{f_{u}^{\mu} | u \in \{1, \ldots, M\}\} \) and the corresponding standard deviation vector \( F_{\sigma} = \{f_{u}^{\sigma} | u \in \{1, \ldots, M\}\} \), where \( f_{u}^{\mu} = 1/|C| \sum_{c \in C} f_{u}(c) \) and \( f_{u}^{\sigma} = \sqrt{1/|C|} \sum_{c \in C} [f_{u}(c) - f_{u}^{\mu}(c)]^{2} \) are obtained. At each \( c \in C \), z score \( S(c) \) is defined as

\[
S(c) = \frac{\|F(c) - F_{\mu}\|^{2}}{\|F_{\sigma}\|^{2}}. \tag{2.3}
\]

where \( F(c) \) corresponds to the feature vector at each \( c \) and \( S(c) \) reflects the degree to which the value of an object deviates from the normal based on a statistical linear model. A predefined threshold \( \theta \) is then used to classify each \( c \in C \) into one class or the other based on whether \( S(c) \geq \theta \).

II.A.3. Nonlinear dimensionality reduction methods

In this work we consider two popular NLDR schemes, locally linear embedding (LLE) (Ref. 30) and GE (Ref. 32) for MRS analysis. We aim to demonstrate the use of NLDR schemes for representation of MRS data results in superior discrimination between CaP and non-CaP spectra compared to the use of linear DR schemes such as PCA.

Graph embedding. The aim of graph embedding \(^3^2\) is to find an embedding vector \( S_{\text{GE}}(c_i), \forall c_i \in C, i \in \{1, \ldots, |C|\} \), such that the relative ordering of the distances between objects in high dimensional space is maximally preserved in the lower dimensional space. Thus, if locations \( c_i, c_j \in C, i, j \in \{1, \ldots, |C|\} \) are adjacent in the high dimensional feature space, then \( ||S_{\text{GE}}(c_i) - S_{\text{GE}}(c_j)||_{2} \) should be small, where \( ||.||_{2} \) represents the Euclidean norm. This will only be true if the distances between all \( c_i, c_j \in C \) are preserved in the low dimensional mapping of the data. To compute the optimal embedding, an adjacency matrix \( W_{\text{GE}} \in \mathbb{R}^{(|C| \times |C|)} \) is first defined as

\[
W_{\text{GE}}(i,j) = e^{-||F(c_i) - F(c_j)||^2}, \quad \forall c_i, c_j \in C, i, j \in \{1, \ldots, |C|\}. \tag{2.4}
\]

\( S_{\text{GE}}(c_i) \) is then obtained via maximization of the following objective function:

\[
E(X_{\text{GE}}) = 2 \gamma \frac{\sum \lambda_{i}}{X_{\text{GE}}(D - W_{\text{GE}})X_{\text{GE}}^{T}} - X_{\text{GE}}D X_{\text{GE}}^{T}, \tag{2.5}
\]

where \( X_{\text{GE}} = [S_{\text{GE}}^{(1)}(c_1), S_{\text{GE}}^{(1)}(c_2), \ldots, S_{\text{GE}}^{(1)}(c_n)], n = |C|, \) and \( \gamma = |C| - 1 \). Additionally, \( D \) is a diagonal matrix where \( \forall c_i \in C, i \in \{1, \ldots, |C|\}, \) the diagonal element is defined as \( D_{i,i} = \sum_{j} W_{\text{GE}}(i,j) \). The embedding space is defined by the eigenvectors corresponding to the smallest \( m \) eigenvalues of \( D - W_{\text{GE}} \) \( X_{\text{GE}} = \lambda D X_{\text{GE}} \). The matrix \( X_{\text{GE}} \in \mathbb{R}^{(|C| \times m)} \) of the first \( m \) eigenvectors of \( X_{\text{GE}} \) is defined as row \( i \) of \( X_{\text{GE}} \).

Locally linear embedding (LLE). LLE (Ref. 30) operates by assuming that objects within a local neighborhood in a high dimensional feature space are linearly related. Consider the set of high dimensional feature vectors \( \mathcal{F} = \{F(c_1), F(c_2), \ldots, F(c_n)\}, \forall c_i \in C, i \in \{1, \ldots, n\} \). LLE aims to map the set \( \mathcal{F} \) to the corresponding set \( \mathcal{S}_{\text{LLE}} = \{S_{\text{LLE}}^{(1)}(c), S_{\text{LLE}}^{(2)}(c), \ldots, S_{\text{LLE}}^{(K)}(c)\} \) of embedding coordinates. Let \( d_{(1)}, \ldots, d_{(K)} \) be the \( K \) nearest neighbors of \( c_i \) and let \( n_{(k)}(c_i) \) be the indices of the location of the \( K \)-nearest neighbors (K-NN) of \( c_i \in C \). The feature vector \( F(c_i) \) and its K-NN’s \( \{F(d_{(1)}), F(d_{(2)}), \ldots, F(d_{(K)})\} \) are assumed to lie on a
patch of the manifold that is locally linear, allowing us to use the Euclidean metric to determine distance between neighbors. Each \( F(c_i) \) can then be approximated by a weighted sum of its \( K \)-NN. The optimal reconstruction weights are given by the sparse matrix \( W_{\text{LLE}} \) (subject to the constraint \( \sum_j W_{\text{LLE}}(i,j)=1 \)) that minimizes

\[
E_1(W_{\text{LLE}}) = \sum_{i=1}^{n} \left\| F(c_i) - \sum_{j=1}^{K} W_{\text{LLE}}(i, j') F(d_{ij}) \right\|_2^2.
\]  

(2.6)

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

\[
E_2(\chi_{\text{LLE}}) = \sum_{i=1}^{n} \left\| S_{\text{LLE}}(c_i) - \sum_{j=1}^{n} W_{\text{LLE}}(i,j) S_{\text{LLE}}(c_j) \right\|_2^2
= \text{tr}(\chi_{\text{LLE}} L \chi_{\text{LLE}}^T),
\]

(2.7)

where \( \chi_{\text{LLE}} = [S_{\text{LLE}}(c_1), S_{\text{LLE}}(c_2), \ldots, S_{\text{LLE}}(c_n)] \), \( L = I - W_{\text{LLE}}/(I - W_{\text{LLE}}) \), and \( I \) is the identity matrix. The minimization of Eq. (2.7) subject to the constraint \( \chi_{\text{LLE}}^T \chi_{\text{LLE}} = I \) (a normalization constraint that prevents the solution \( \chi_{\text{LLE}} = 0 \)) is an eigenvalue problem whose solutions are the eigenvectors of the Laplacian matrix \( L \). Since the rank of \( L \) is \( n-1 \), the first eigenvector is ignored and the second smallest eigenvector represents the best one-dimensional projection of all the samples. The best two-dimensional projection is given by the eigenvectors with the second and third smallest eigenvalues, and so forth.

II.B. Replicated \( k \)-means clustering in the reduced feature space

For the DR schemes, \( \phi \in \{ \text{PCA}, \text{LLE}, \text{GE}\} \), unsupervised replicated clustering is used to classify all objects \( c \in C \) into one of the \( k \) classes based on \( S^\phi(c) \), the low dimensional representation of \( F(c) \). Replicated clustering is a variant of the popular \( k \)-means\(^{34} \) clustering scheme. The \( k \)-means algorithm is initialized by randomly partitioning the data into \( k \) clusters and computing the cluster center for each partition. The distance of each point from each of the \( k \) centroids is computed and each object is reassigned to the closest cluster centroid to minimize the intraclass variance. This random initialization may lead to local minima leading, in turn, to different clustering results. The motivation behind replicated clustering is to make the final aggregation results from \( k \)-means more deterministic. In replicated clustering, multiple weak clusterings of the data are generated. The optimal clustering solution is then chosen among the various weak clusterings as the one with the least intracluster variance. Below we briefly describe the various steps involved in this algorithm.

Step 1. At each of \( T \) iterations, \( k \)-means is applied to cluster all objects \( c \in C \) into one of the \( k \) classes \( V_1^t, V_2^t, \ldots, V_k^t \),

\[
\text{MRS metavoxel } c \in C
\]

\[
\text{MRI voxels } g \in G
\]

Fig. 3. Illustration of the spatial relationship between MRS metavoxels \( c \in C \) and T2-w MRI voxels \( g \in G \). The spectral grid \( C \) comprising of 28 metavoxels has been overlaid on a T2-w MRI prostate slice and is shown in white. Note the region outlined in red on \( C \) corresponds to the area occupied by a metavoxel, but may contain multiple MRI voxels (highlighted in red).

\[
t \in \{1, \ldots, T\}, \text{ where each } c \in C \text{ is characterized by a high dimensional feature vector } F(c). \text{ For each } c \in V_1^t, q \in \{1, 2, \ldots, k\}, t \in \{1, 2, \ldots, T\} \text{ the centroid of each cluster is determined as }
\]

\[
F_q^t = \frac{1}{|V_q^t|} \sum_{c \in V_q^t} F(c).
\]

(2.8)

Step 2. At each iteration \( t \in \{1, \ldots, T\} \), the average Euclidean distance between each \( F(c) \in V_q^t \) and corresponding cluster center \( F_q^t \), \( t \in \{1, \ldots, T\}, q \in \{1, 2, \ldots, k\} \), is then determined as

\[
d_q^t = \frac{1}{|V_q^t|} \sum_{c \in V_q^t} ||F(c) - F_q^t||.
\]

(2.9)

The average intracluster distance over all \( k \) clusters is then obtained as

\[
\mu_q^t = \frac{1}{k} \sum_{q} d_q^t.
\]

(2.10)

Step 3. Finally, the clusterings \( \hat{V}_q, q \in \{1, 2, \ldots, k\} \) within a specific iteration \( t \in \{1, \ldots, T\} \) are identified as the stable clustering result for which \( \mu_q^t \) is minimum over all \( t \).

Note that replicated clustering identifies stable clusterings as those that minimize intraclass variance. Note further that while we are not explicitly seeking to increase intercluster distance, our empirical results suggest that replicated \( k \)-means clustering tends to also push the cluster centers farther apart.

III. METHODOLOGY

In Sec. III A we provide a brief description of the notation employed in this paper. In Sec. III B we provide a brief description of the data sets considered in this study. Details concerning determination of the approximate ground truth for spatial extent of CaP on MRI are provided in Sec. III C. Methodological details regarding the approximate ground truth for CaP detection on MRS are detailed in Secs. III D and III E, respectively.

III.A. Notation

We represent the 3D prostate T2-w MRI scene by \( G = (G, f) \), where \( G \) is a 3D grid of voxels \( g \in G \) and \( f(g) \) is a function that assigns an intensity value to every \( g \in G \). We
also define a spectral scene \( C = (C, F) \) where \( C \) is a 3D grid of metavoxels, \( c \in C \). Each metavoxel \( c \) is associated with a corresponding \( M \)-dimensional spectral vector \( F(c) = [f_1(c), \ldots, f_M(c)] \), where \( f_j(c) \) represents the MRS signal intensity at each \( c \) along the frequency domain. Figure 3 shows the spatial relationship between the MR spectral metavoxel \( c \in C \) and T2-w MRI voxel \( g \in G \). Note that the distance between any two adjacent metavoxels \( c_i, c_j \in C \), \( \|c_i - c_j\|_2 \), where \( \| \cdot \|_2 \) denotes the \( L_2 \) norm, \( i, j \in \{1, \ldots, |C|\} \), and \( |C| \) is the cardinality of \( C \), is roughly 13 times the distance between any two adjacent MRI voxels \( g_i, g_j \in G \), where \( i, j \in \{1, \ldots, |G|\} \). A list of commonly used notations and symbols in this paper is given in Table I.

### III.B. MRSI data description

A total of 18 deidentified and anonymized prostate cancer MRSI/MRI datasets from the ACRIN trial was chosen randomly for this pilot study. All exams were performed on a 1.5 T magnet (GE Medical Systems, Milwaukee, WI). Oblique axial T2-w spectra (TR 4000–6000 ms, TE 90–120 ms, slice thickness of 3 mm, acquisition matrix of 256 × 192, and FOV of 12–14 cm) were first obtained. Using the PROSE (Prostate spectroscopy and imaging examination) software package (www.getecare.com/user/mr/applications/products/prose.html), MRSI (TR 1000 ms, TE 130 ms, spectral width of 1000 Hz, and number of points is 512) was then prescribed from the oblique axial T2-w images, with outer voxel suppression and oblique suppression planes manually set to exclude the majority of fat about the prostate capsule. A FOV of 11 cm was used for the 16 × 16 MRSI grid (voxel dimension of 6.75 × 6.75 × 3 mm³). MRSI was obtained as a PRESS (point resolved spectroscopy sequence) sequence. The MRS spectral grid was contained in DICOM image sets, from which the 16 × 16 grid containing 256 complex spectra per slice was obtained using IDL 6.4 (ITT Visual Information Systems).

### III.C. Determining approximate ground truth for spatial extent of CaP on MRI-MRS

Following radical prostatectomy, the gland was fixed in formalin and sectioned per institutional routine (whole mount or standard sections). Sections were then embedded in formalin, and duplicate slides for each block were prepared for central review. Slides were stained with Hematoxylin and Eosin and reviewed by a single central pathologist for areas of cancer. MRI-pathology correlation was established through a joint review session of trial imagers and pathologists, who determined slice-by-slice histology-MRI concordance. In order to ensure uniform spectral and histologic sextant assignment, sextant boundaries (i.e., between apex and midgland, and between midgland and base) were based on the MRI slice assignments that had previously been determined by each site radiologist. Using these sextant boundaries, and the best approximation of MRSI-histologic concordance, the presence and the diameter of CaP in each sextant were established with maximum tumor diameter, denoted as \( R \), recorded for all positive sextants.

### III.C.1. Potential cancer space

In order to quantitatively evaluate CAD performance in terms of performance metrics such as sensitivity and specificity, the precise spatial location of the target class within \( C \) is required. Unfortunately, this information was not readily available for the studies considered in this project. Hence, we define a probabilistic ground truth for CaP that involves first defining a potential cancer space representing a spatial region \( C_p \) within \( C \) within which the tumor is embedded. To appreciate the need for \( C_p \), let us assume the ideal case scenario (Fig. 4) where the precise spatial location of CaP is known \( a \ priory \) and is denoted by the set \( C_i \). If \( C_{i,j} \) denotes the set of spatial locations corresponding to CaP identified by the CAD system, the true positive (TP) area could be calculated as \( |C_i \cap C_{i,j}| \). Similarly, the false positive (FP) area for the CAD system is \( |C_{i,j} - (C_i \cap C_{i,j})| \) and false negative (FN) area is \( |C_i - (C_i \cap C_{i,j})| \). For the problem we are considering,
$C_i$ is unavailable and hence the need for $C_p$ which is defined as the set of spatial locations within which a total number of $N_g$ CaP locations are contained. The true number of metavoxels $c$ within $C_p$ that represent CaP can be calculated as

$$N_g = K_s \times \left( \frac{(R^2)}{\Delta X \Delta Y} \right),$$

where $K_s$ represents the number of contiguous MR sections containing CaP, $[\cdot]$ refers to the ceiling operation, and $\Delta X$ and $\Delta Y$ refer to the size of the metavoxel $c$ in the $X$ and $Y$ dimensions. Thus for a MRS scene $C$, with known cancer in left midgland (LM), the prostate being contained in a $3 \times 6$ grid and the prostate midgland region extending over two contiguous slices, the total number of CaP metavoxels $|C_p|$ is 18 $(3 \times 3 \times 2)$. The $3 \times 6$ prostate grid is divided into two equal right and left halves. Given that the tumor has a maximum diameter of 13.75 mm in LM, with $\Delta X, \Delta Y=6.875$, $N_g=8$ metavoxels corresponding to CaP within $C_p$. Both $C_p$ and $N_g$ are integral to defining probabilistic estimates of CAD sensitivity, specificity, and positive predictive value (PPV), details of which are provided in Sec. IV.

### III.C.2. Expert annotations

We also endeavored to obtain a more precise estimate of CaP ground truth for a subset of the studies via expert annotation of individual MR spectra. For 6 of the 18 ACRIN studies, an expert radiologist laboriously annotated each individual spectra with labels 1–5; with 1 and 2 being spectra that appeared to be normal and spectra 3–5 being suspicious appearing. For this expert annotated dataset, our CAD system attempted to discriminate all spectra as being either normal or suspicious appearing (i.e., distinguishing between spectra labeled as 1/2 and those labeled as 3/4/5).

### III.D. Stage 1: Localization of prostate using hierarchical spectral clustering of MRS

For the FOV for studies considered in this work, a majority of the spectra lay outside the prostate. The motivation behind stage I of our algorithm is to exploit the differences in spectral characteristics of spectra inside and outside the prostate to hone in the prostate ROI automatically. This is done by identifying and eliminating the dominant cluster at every iteration until a cluster size threshold is attained. The nonlinear DR scheme graph embedding is employed to project all spectra into a reduced dimensional embedding $S^{\theta}(c)$, followed by replicated k-means clustering to aggregate all $c \in C$ into two clusters $V^1, V^2$ corresponding to prostatic and extracapsular spectral classes. At each iteration $t \in \{1, \ldots, T\}$, a subset of voxels $C_t$ in $C$ is obtained by eliminating the noninformative extracapsular spectra identified as the dominant cluster ($V^{dom}$). The approximate number of prostate spectra ($\Theta$) of a MRS grid is learned during the offline training phase. The automatic cascaded scheme stops when the number of remaining spectra in the MRS grid is approximately equal to $\Theta$. The result of the HierarclustMRS algorithm is a spectral grid ($\tilde{C}_T$) containing all the prostate spectra.

#### Algorithm HierarclustMRS

**Input**: $F(c)$ for all $c \in C, \Theta, C$.  
**Output**: $\tilde{C}_T$

```plaintext
begin
0. Initialize $\tilde{C}_0 = C, t = 0$
1. while $|\tilde{C}_t| > \Theta$ do
2. Apply Graph Embedding (Ref. 32), to $F(c)$, for all $c \in \tilde{C}_t$ to obtain $S^{\Theta}(c)$
3. Apply replicated k-means clustering on $S^{\Theta}(c)$ to obtain two stable clusters $V_1, V_2$
4. Identify larger cluster $V_1^{dom} = \arg \max_{c \in V_1}(|V_1|)$, where $w \in \{1, 2\}$
5. Create set $\tilde{C}_t+1 \subset \tilde{C}_t$ by eliminating all $c \in V_1^{dom}$ from $\tilde{C}_t$
6. $t = t + 1$
7. endwhile
8. $\tilde{C}_T = \tilde{C}_T$
end
```

Note that in general the algorithm is terminated when the total number of spectra is marginally greater than (or equal to) $\Theta$, which usually occurs within two to three iterations.

### III.E. Stage 2: CaP identification on MRS via NLDR

Following stage 1, we attempt to apply more sophisticated analysis to the spectra in $\tilde{C}_T$ to be able to discriminate between different tissue classes in the prostate. Apart from NLDR schemes (LLE and graph embedding) that were considered, two other feature extraction schemes ($z$ score and PCA) were also evaluated in terms of their ability to discriminate between different classes of the prostate spectra (abnormal appearing, normal appearing and indeterminate). Following feature extraction, replicated k-means clustering for the DR schemes (LLE, graph embedding, and PCA) and thresholding for $z$ score was applied to obtain hard classification of the spectra into CaP and non-CaP categories. For $z$ score evaluation, a predefined threshold $\theta$ is used to classify each $c \in \tilde{C}_T$ as abnormal or normal appearing based on whether $S^\theta(c) > \theta$ or $S^\theta(c) < \theta$. Three commonly employed DR methods (explained previously in Sec. III) are applied to the MR spectra in $\tilde{C}_T$ so that for any $c \in \tilde{C}_T$, the high dimensional ambient feature vector $F(c)$ is mapped to $S^\theta(c)$, where
\( \phi \in \{ \text{PCA}, \text{LLE}, \text{GE} \} \). Replicated clustering is then employed to cluster each \( S^\phi(c), \forall c \in \tilde{C}_T \), into one of the three possible classes, \( \tilde{\phi}^1, \tilde{\phi}^2, \tilde{\phi}^3 \) corresponding to abnormal appearing, normal appearing, or an indeterminate class, with the intermediate class corresponding to undefined, equivocal, or noisy spectra. For the manually annotated datasets, replicated clustering was used to identify all analyzable spectra as being either normal or suspicious appearing (two clusters).

**IV. EVALUATION METHODS**

**IV.A. Identification of cancer cluster**

As mentioned previously in Sec. III C 1, a potential CaP space \( C_p \) is defined within which the number of CaP locations \( N_g \) is determined. Following replicated clustering of \( S^\phi(c), c \in \tilde{C}_T \), we need to identify which of the \( k \) clusters \( \tilde{\phi}^1, \tilde{\phi}^2, \ldots, \tilde{\phi}^k \) \( \in \tilde{C}_T \) correspond to the CaP class. We represent the corresponding sets of spatial locations for clusters \( \tilde{\phi}^1, \tilde{\phi}^2, \ldots, \tilde{\phi}^k \) as \( C_1^a, C_2^a, \ldots, C_k^a \) respectively. With respect to \( C_p \) (as illustrated in Fig. 5), some part of \( C_g^a \) may be within \((C_1^a, q \in \{1,2,\ldots,k\})\) or outside \((C_1^a, q \in \{1,2,\ldots,k\})\) \( C_p \). Thus \( C_g^a = C_1^a \cup C_2^a, q \in \{1,2,\ldots,k\} \). The TP, FP, true negative (TN), and FN ratios associated with each class \( C_g^a, q \in \{1,2,\ldots,k\} \) with respect to \( C_p \) and \( N_g \) are then obtained. The following heuristic algorithm is then used to identify which of \( C_g^a, q \in \{1,2,\ldots,k\} \) represents the CaP class \( (C_p^a) \).

**Algorithm IdentifyCaPClass**

**Input:** \( C_g^a, q \in \{1,2,\ldots,k\}, C_p, N_g, C \).

**Output:** \( C_p^a \)

1. \( q = 1 \) to \( k \) do;
2. \( C_g^a = C_g^a \cap C_p \);
3. \( C_g^a = C_g^a - C_g^a \);
4. if \( |C_g^a| > N_g \) then
   1. \( TP := N_g, FN := |C_g^a| - N_g, TN := |C_g^a| - N_g \);
   2. else \( TP := |C_g^a|, FN := N_g - |C_g^a|, TN := |C_g^a| - N_g \);
5. end;
6. end;
7. \( q^* = \arg \max \left( \frac{TP}{TN} \right) \);
8. \( C_p^a = C_g^a \);

Thus Algorithm IdentifyCaPClass identifies the CaP cluster as the one that maximizes true positive fraction while simultaneously minimizing the false positive fraction.

**V. RESULTS**

**V.A. Qualitative results**

**V.A.1. Stage 1: Qualitative evaluation of the hierarchical clustering scheme**

Figure 6 shows the qualitative results of the hierarchical cascade scheme for distinguishing prostatic from extracapsular spectra. Figure 6(a) represents spatial maps of the spectral grid \( \tilde{C}_0 \) (16 x 16 spectral voxels) superimposed on the corresponding T2-w MRI scene for one patient study. Every ca \( c \in \tilde{C}_0 \) in Fig. 6(a) is assigned one of two colors (blue and red) corresponding to spectra identified by the algorithm as prostatic or extracapsular. Note that the dominant cluster [spatial locations in red in Fig. 6(a)] has been eliminated in the second iteration \([\tilde{C}_1 \ (16 \times 8 \ spectral \ voxels)]\) [Fig. 6(b)]. The final spectral grid \([\tilde{C}_2 \ in \ Fig. \ 6(c)]\) is obtained after elimination of extracapsular spectra (red locations) during the third iteration of the cascade. Figures 6(d)–(f) represent the embedding plots [where each original spectral vector \( F(c), c \in C \) is plotted in 3D eigenvector space using the three dominant embedding values as coordinates] from \( \tilde{C}_0 \) (16 x 16 spectral voxels) (d) to \( \tilde{C}_2 \ (7 \times 4 \ spectral \ voxels) \) (f) for one study at three different levels of the cascade. Note that at the end of the third iteration, the prostate ROI has been accurately identified and the spectral grid accurately overlaid on the prostate. Further note that in Figs. 6(a)–6(c), the spectral grid with the pronounced boundary indicates the ROI during the current iteration.
V.A.2. Stage 2: Evaluation of feature extraction schemes for CaP detection

The identification of the prostate grid in stage 1 allows for the resolvability of the three MR spectral classes (abnormal appearing, normal appearing, and indeterminate). The differences between these three spectral classes within the prostate spectra cluster at the higher levels in the cascade \((\tilde{C}_1, \tilde{C}_2, \tilde{C}_3)\) become discriminable only after the removal of extracapsular spectra. Note that replicated clustering was used to identify only two clusters (suspicious and normal appearing spectra) for the six manually annotated studies. Figure 7 shows the qualitative results of the four feature extraction schemes employed in this work for CaP detection for two different patient studies, each row in Fig. 7 corresponding to a different study. The three colors assigned to the spectral voxels in Fig. 7 correspond to the three clusters obtained via replicated clustering on the reduced dimensional spectra \(S^{\phi}(c), \phi \in \{\text{PCA, LLE, GE}\}\) for \(c \in C\). For the \(z\) score scheme, each metavoxel was classified as belonging to one of the two classes [red and blue in Figs. 7(a) and 7(d)], reflecting abnormal and normal appearing spectra. The white box superposed in Figs. 7(a)–7(f) shows the potential cancer space \((C_P)\) for corresponding slices. In each of Figs. 7(a)–7(f), the red cluster was identified as comprising abnormal using the IdentifyCaPCluster algorithm (sec. IV) and following feature extraction and replicated clustering. Figures 7(a) and 7(d) show the results obtained via \(z\) score; while Figs. 7(b) and 7(e) show the results of PCA in identifying abnormal appearing, normal appearing, and indeterminate spectra. Figures 7(c) and 7(f) show the corresponding results for graph embedding and LLE, respectively. For the first study (first row in Fig. 7), graph embedding [Fig. 7(c)] appears to yield a near perfect CaP detection in terms of sensitivity and specificity as only CaP voxels are identified within the white cancer grid \((C_P)\). The corresponding results for \(z\) score [Fig. 7(a)] and PCA [Fig. 7(b)] both yield poor detection sensitivity and specificity. Similarly, for the second study shown in Fig. 7, LLE [Fig. 7(f)] appears to yield higher CaP detection sensitivity and specificity compared to \(z\) score [Fig. 7(d)], and PCA [Fig. 7(e)]. Figure 8 shows an example of the MR spectral grid with classification labels obtained from graph embedding and replicated clustering plotted back on the individual spectra; the spectra in red corresponding to those identified as abnormal appearing, blue spectra corresponding to normal appearing, and green spectra corresponding to indeterminate.

V.B. Quantitative results

V.B.1. Quantitative evaluation of stage 1: Hierarchical clustering

At the end of stage 1, the largest rectangular box \(\tilde{C}_T\) that contains all prostate spectra is then overlaid on the T2-w image. Note that in quantitative evaluation of stage 1, the precise spatial extent of the prostate is all that is needed. This ground truth is ascertained by manual placement of a spectral grid \((\tilde{C}_g)\) on the prostate by an expert radiologist. Table II

Fig. 6. Spectral grids for a single 2D slice of a T2-w MRI scene for a patient at (a) the first cascade level \(\tilde{C}_0\), (b) second cascade level \(\tilde{C}_1\), and (c) third cascade level \(\tilde{C}_2\). Note that the size of the grid reduces from \(16 \times 16\) metavoxels in (a) to \(7 \times 4\) in (c) by elimination of extracapsular spectra in the dominant cluster (red). The corresponding clustered embedding plots at each of the cascade levels are also shown in (d)–(f), corresponding, in turn, to the metavoxel grids shown in (a)–(c).
Defined during an offline training phase. For each $S$ corresponding values for CaP. The threshold $C_{\gamma}$ was determined as belonging to CaP if contained via annotation by an expert radiologist. At each value of $C_{\gamma}$, where $C_{\gamma}$ was then defined as the operating point $z$ score-based classification of each $C_{\gamma}$, was assumed. In order to evaluate the parameter sensitivity of different DR methods, the robustness of the DR methods over different values of $\kappa$ and $v$ was quantitatively evaluated. LLE was evaluated by varying $\kappa \in \{6, 7, \ldots, 15\}$ and $v \in \{3, 4, \ldots, 10\}$, a total of 80 different combinations of parameter values. PCA and graph embedding were evaluated for eight different values of $v \in \{3, 4, \ldots, 10\}$.

For each of $\phi \in \{\text{LLE, GE, PCA}\}$, threefold and fivefold cross validation averaged over 20 iterations were also performed to obtain average sensitivity and specificity in terms of CaP detection performance for all 18 studies. Threefold cross validation was performed by randomly choosing three datasets and calculating average CaP detection sensitivity and specificity across three studies on the 80-dimensional parameter space. The parameter set $(v_{\text{max}}, \kappa_{\text{max}}$ for LLE) with maximum sensitivity and specificity in this space were then identified as optimal values and used for the CAD scheme on the remaining 15 studies. The average CaP detection sensitivity and specificity on these 15 studies was then recorded. On the next trial, 3 random training studies from 18 were again selected and used to optimize the parameters and evaluation again was done on the remain-
Fig. 8. MRS spectral grid plotted with the classification labels (three colors correspond to three different clusters) obtained from graph embedding and replicated clustering. The red spectra correspond to those identified as abnormal appearing, the blue correspond to normal appearing, and the green correspond to indeterminate spectra (possibly noise in this case).

ing 15 studies. This entire process was repeated for a total of 20 times. The mean $\mu^\phi$ and standard deviation $\sigma^\phi$ in CaP detection sensitivity and specificity across these 20 iterations were reported in Table III(a) for $\phi \in \{\text{LLE, GE, PCA}\}$. A similar routine was employed when performing fivefold cross validation. The corresponding results are reported in Table III(b). All NLDR schemes employed in this work were found to have higher sensitivity of CaP detection compared to PCA; graph embedding performing the best with a sensitivity of almost 90%. In Figs. 9 and 10 barplots are shown representing CaP detection sensitivity and specificity ($\mu^\phi, \sigma^\phi$) for each of the 18 studies considered in this work via threefold and fivefold cross validation, respectively, over 20 iterations for GE. The confidence estimates (\(\eta\)) associated with the sensitivity and specificity measurements of each study are also shown. Note that while \(\eta\) is low for a majority of the studies, for studies 11 and 18, the confidence associated with CaP detection sensitivity and specificity for graph embedding were almost 90% and 80%, respectively.

Evaluation of all feature extraction methods against expert annotated spectra. Table IV shows the average sensitivity and specificity results obtained for $\phi \in \{\text{LLE, GE, PCA}\}$

<table>
<thead>
<tr>
<th>Method</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>89.33 ± 1.87</td>
<td>79.79 ± 2.24</td>
</tr>
<tr>
<td>LLE</td>
<td>83.70 ± 4.50</td>
<td>81.04 ± 4.60</td>
</tr>
<tr>
<td>PCA</td>
<td>78.53 ± 3.20</td>
<td>83.97 ± 2.81</td>
</tr>
</tbody>
</table>

(a)

(b)

<table>
<thead>
<tr>
<th>Method</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>89.09 ± 2.10</td>
<td>79.22 ± 2.10</td>
</tr>
<tr>
<td>LLE</td>
<td>84.93 ± 4.42</td>
<td>81.32 ± 4.05</td>
</tr>
<tr>
<td>PCA</td>
<td>78.78 ± 4.31</td>
<td>84.11 ± 2.46</td>
</tr>
</tbody>
</table>

Table III. This shows the average and standard deviation in CaP detection sensitivity and specificity for different feature extraction methods over $\kappa \in \{5, \ldots, 15\}$ and $\upsilon \in \{3, \ldots, 10\}$ for 18 different studies via (a) threefold cross validation, and (b) fivefold cross validation.
VI. CONCLUDING REMARKS

In this paper we have presented a novel application of nonlinear dimensionality reduction and hierarchical clustering for automated identification of (a) the prostate ROI based on the classification of MR spectral data alone and (b) suspicious appearing spectra within the prostate ROI. A total of 18 MRI/MRS studies from the ACRIN trial was considered for the evaluation of four different feature extraction algorithms (PCA, \( z \) score, LLE, GE) in conjunction with replicated clustering in terms of their ability to identify suspicious appearing spectra. Owing to the fact that only limited knowledge regarding precise spatial extent of CaP was available for the studies considered in this work, we defined a probabilistic ground truth estimate for CaP and a confidence coefficient to assess the degree of certainty associated with the CAD performance measures reported. The high confidence estimates associated with two of the studies (study 11 and 18) seem to suggest that the consistently high CaP detection sensitivity and specificity measurements for the other studies are not erroneous. For a subset (6) of the 18 studies, a radiologist laboriously annotated each of the spectra according to the five point scale\(^{36} \) by an expert radiologist. For all methods, we made the assumption that spectra labeled as 1/2 were all normal appearing and the spectra labeled 3/4/5 were suspicious appearing. For each of \( \phi \in \{ \text{LLE, GE, PCA, } z \} \), the IdentifyCaPCluster algorithm was employed to identify the cluster most likely to correspond to CaP spectra. Since we only had six studies available, the mean and standard deviation in CaP detection sensitivity and specificity across all the expert annotated datasets were determined (based on previously determined \( u_{\text{max}} \) and \( k_{\text{max}} \)) and shown in Table IV.

**Table IV.** This shows the average and standard deviation in CaP detection sensitivity and specificity for different feature extraction methods obtained via comparison with expert annotations for six ACRIN studies.

<table>
<thead>
<tr>
<th>Method</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>79.74 ± 13.18</td>
<td>65.49 ± 7.58</td>
</tr>
<tr>
<td>LLE</td>
<td>81.36 ± 5.69</td>
<td>64.71 ± 9.84</td>
</tr>
<tr>
<td>PCA</td>
<td>76.68 ± 14.07</td>
<td>48.26 ± 12.95</td>
</tr>
<tr>
<td>( z ) score</td>
<td>62.70 ± 18.67</td>
<td>50.73 ± 7.39</td>
</tr>
</tbody>
</table>

ACKNOWLEDGMENTS

This work was made possible via grants from the Wallace H. Coulter Foundation, New Jersey Commission on Cancer Research, National Cancer Institute (Grant Nos.
R01CA136535-01, ARRA-NCI-3 R21CA127186–02S1, R21CA127186–01, and R03CA128081-01), the Society for Imaging Informatics in Medicine (SIIM), The Cancer Institute of New Jersey, and the Life Science Commercialization Award from Rutgers University.

A1 Author to whom correspondence should be addressed: 599 Taylor Road, Piscataway, NJ 08854. Telephone: 732-445-4500 (ext. 6213). Fax: 732-445-3753. Electronic mail: anantm@rci.rutgers.edu


