Improving supervised classification accuracy using non-rigid multimodal image registration: Detecting Prostate Cancer

Jonathan Chappelow[a], Satish Viswanath[a], James Monaco[a], Mark Rosen[b], John Tomaszewski[c], Michael Feldman[c], Anant Madabhushi[a]

[a]Rutgers University Department of Biomedical Engineering, 599 Taylor Road, Piscataway, NJ, USA 08854

[b]Department of Radiology, [c]Department of Pathology, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA, USA 19104

ABSTRACT

Computer-aided diagnosis (CAD) systems for the detection of cancer in medical images require precise labeling of training data. For magnetic resonance (MR) imaging (MRI) of the prostate, training labels define the spatial extent of prostate cancer (CaP); the most common source for these labels is expert segmentations. When ancillary data such as whole mount histology (WMH) sections, which provide the gold standard for cancer ground truth, are available, the manual labeling of CaP can be improved by referencing WMH. However, manual segmentation is error prone, time consuming and not reproducible. Therefore, we present the use of multimodal image registration to automatically and accurately transcribe CaP from histology onto MRI following alignment of the two modalities, in order to improve the quality of training data and hence classifier performance. We quantitatively demonstrate the superiority of this registration-based methodology by comparing its results to the manual CaP annotation of expert radiologists. Five supervised CAD classifiers were trained using the labels for CaP extent on MRI obtained by the expert and 4 different registration techniques. Two of the registration methods were affine schemes; one based on maximization of mutual information (MI) and the other method that we previously developed, Combined Feature Ensemble Mutual Information (COFEMI), which incorporates high-order statistical features for robust multimodal registration. Two non-rigid schemes were obtained by succeeding the two affine registration methods with an elastic deformation step using thin-plate splines (TPS). In the absence of definitive ground truth for CaP extent on MRI, classifier accuracy was evaluated against 7 ground truth surrogates obtained by different combinations of the expert and registration segmentations. For 26 multimodal MRI-WMH image pairs, all four registration methods produced a higher area under the receiver operating characteristic curve compared to that obtained from expert annotation. These results suggest that in the presence of additional multimodal image information one can obtain more accurate object annotations than achievable via expert delineation despite vast differences between modalities that hinder image registration.

Keywords: registration, prostate cancer, CAD, dimensionality reduction, mutual information, thin plate splines, non-rigid, COFEMI, histology, MRI, multimodal, independent component analysis, Bayesian classifier

1. INTRODUCTION

Supervised classifiers are the central component of computer-aided diagnosis (CAD) systems for medical imagery, in which the goal is to classify observations (e.g. pixels) as belonging to a particular class (e.g. cancer or non-cancer) based on prior observations. A crucial step in the implementation of such classifiers is obtaining object labels that accurately represent the targeted tissue class. The most common way of obtaining such class labels involves manual labeling by an expert (e.g. radiologist). Although, even with trained experts, manual annotations are time consuming, variable, and error prone.[1] The adverse effect of any mislabeled training instances on supervised classifier performance has been repeatedly observed[2–4] in terms of accuracy and robustness. As such, avoidance of incorrectly labeled instances is of utmost importance in building an effective supervised CAD classifier.

Contact: Anant Madabhushi: E-mail: anantm@rci.rutgers.edu, Telephone: 1 732 445 4500 x6213

Medical Imaging 2008: Computer-Aided Diagnosis, edited by Maryellen L. Giger, Nico Karssemeijer
Proc. of SPIE Vol. 6915, 69150V, (2008) · 1605-7422/08/$18 · doi: 10.1117/12.770703

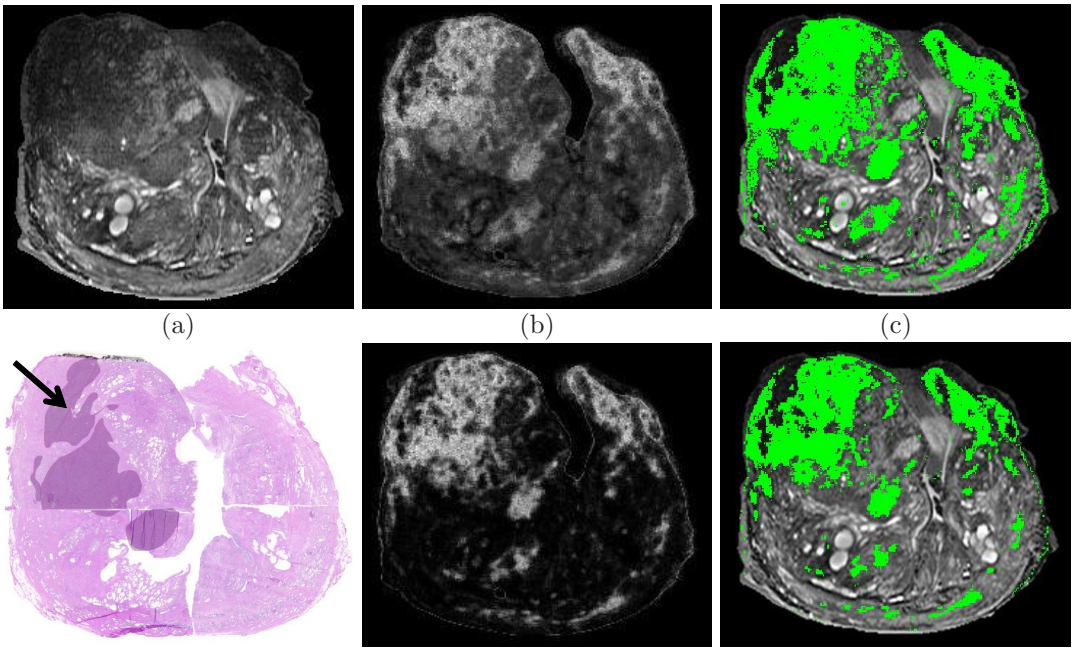


Figure 1. (a) MRI slice of a prostatectomy specimen with cancer present. (b) Likelihood image generated by a CAD system applied to (a), and (c) the prediction for cancer presence obtained by thresholding (b) at 50% confidence level and superposing on (a). (d) The corresponding histology slice shown labeled for cancer extent by H&E stain (black arrow). (e) Likelihood image calculated by CAD after removal of erroneous training instances, and (f) the resulting binary prediction.

Consider the MRI image of a prostate gland is shown in Figure 1(a) from which a CAD system[1] driven by a Bayesian classifier calculates the likelihood that each pixel belongs to the cancer class (Figure 1(b)) and obtaining a binary prediction for cancer presence (Figure 1(c)). Figure 1(e) shows the likelihood image obtained by CAD after removal of erroneous training instances, indicating that removal of even a small portion of mislabeled data resulted in a significant improvement in classification accuracy as can be discerned in the increase in contrast in the posterior likelihood scenes, indicating enhanced separation between classes (compare Figures 1(e) and (b)). Figure 1(d) shows a reconstructed WMH section of the prostate obtained via radical prostatectomy which corresponds to the MRI slice in Figure 1(a).

Delineating regions of CaP on MRI is difficult even for radiologists; fortunately, images from multiple modalities can be utilized to assist in labeling the CAD targeted modality. In this study, histology provides a modality with readily available ground truth for CaP extent, found with the aid of hematoxylin and eosin (H&E) staining. For instance, manual mapping of histological ground truth onto MRI using anatomical features as visual cues has been previously demonstrated in [1], where CaP extent on MRI was obtained by examination of WMH and radiological sections by a pathologist and a radiologist simultaneously. For example, the ground truth for CaP extent provided by the WMH section in Figure 1(d) is manually mapped by the simultaneous efforts of both experts onto the corresponding MRI in Figure 1(a). Despite possession of empirical domain knowledge by the experts performing the mapping, this process is also time consuming and subject to observer bias. Note that this task is akin to manual registration or alignment of the two modalities, whereby a mapping between corresponding pixels is determined. Thus, it is clear that in order to accurately establish training instances using multimodal imagery, it is necessary to achieve accurate registration.

In this view of the manual mapping process, we present an automated labeling scheme whereby the histological ground truth is automatically mapped directly to MR from aligned WMH sections via image registration techniques. The optimal alignment between two images $\mathcal{A}$ and $\mathcal{B}$ is obtained when the pixels representing the same material points on the anatomical object or region of interest (C_a in $\mathcal{A}$, C_b in $\mathcal{B}$) are mapped into the same spatial coordinates ($C_b = T(C_a)$). The result of successful image registration is a pair of images $\hat{\mathcal{A}}$ and $\mathcal{B}$ where $\hat{\mathcal{A}}$ represents $\mathcal{A}$ subjected to the determined transformation T, and pixels C_a in $\hat{\mathcal{A}}$ and C_b in $\mathcal{B}$ represent the same underlying anatomical features of the region of interest. If we consider the set of pixels that represents the

CaP extent on histology, it is clear that determining the transform T, which specifies the mapping of pixels on WMH to MRI, can provide a precise ground truth estimate for CaP on MRI.

In summary, we present a method for automated training of a supervised classifier by establishing accurate training instances via registration of multimodal images. We demonstrate the method for a specific CAD application where the objective is to automatically identify CaP on high resolution *ex vivo* MRI. We further quantitatively demonstrate that the object class labels established by registration are more reliable compared to expert annotations even when (1) considerable image differences exist between modalities and (2) the manual segmentations are meticulously obtained by multiple experts analyzing both modalities simultaneously. The principal contributions of this work are to demonstrate,

- a novel method for automatically training a supervised classifier when multimodal data is present,
- the ability of multimodal registration methods to accurately annotate object labels with significantly greater accuracy compared to expert delineations in spite of differences between disparate modalities (e.g. histology and MRI) which generally hinder successful registration, and
- the potential use of classifier performance as a benchmark for evaluating registration performance.

The rest of this paper is organized as follows. In Section 2 we outline the design of the supervised classifier training technique and its evaluation. In Section 3 we describe the registration methodologies, the classifier and the automated training method, and the cross-validation procedure used to compare the accuracy of multiple classifiers. In Section 6 we present the results of each supervised classifier. Concluding remarks and future research directions are presented in Section 7.

2. OVERVIEW OF EXPERIMENTAL DESIGN

There has been relatively little work in the area of prostate registration, less so in the context of multimodal registration. Registration of prostate histology with ultrasound images has been addressed recently by Taylor[5] who utilized surface matching, Moskalik[6] who fit the prostate boundary with an ellipsoid to find an affine mapping and Porter[7] who used blood vessels as fiducial markers. Lee[8] performed non-rigid registration of CT/MRI with histological sections for a visualization system. In this study, we consider two rigid and two non-rigid multimodal registration methods for aligning prostate histology with MRI. Finding a transformation T which can bring the pixels representing the prostate on histology and MRI into spatial correspondence is hindered by:

1. Dissimilarity of the modalities and the structural characteristics captured by the modalities.
2. The complicated nature of the deformations to which the anatomical object is subjected.

Namely, the processes of quartering, sectioning and fixation of the gland in formalin for acquisition of digitized histological images subject the tissue to irregular and extensive deformation and alter the underlying tissue.

Most current robust non-linear multimodal image registration techniques operate by initially determining an affine or rigid alignment as a precursor to an elastic deformation step.[9–11] An accurate initial global alignment is often important for the success of the elastic alignment step, primarily because physically realistic deformations are most accurately modeled when required deformations are small. Thus, in [12] we presented our registration technique, termed combined feature ensemble mutual information or COFEMI, with emphasis on obtaining an accurate global affine precursor to elastic deformation. In this study, we demonstrate two similarity measures in the optimization of an affine transformation: MI and COFEMI, both followed by elastic transformation. To compare the influence of the affine alignments on the final alignment, we utilize a control point driven thin plate splines (TPS) deformation model to elastically transform histology onto MRI. Thus, the design of our supervised classifier training method and its evaluation against expert-based training are outlined in Figure 2 using the notation defined in Section 3.1 for a set of multimodal images:

1. Automated affine registration of MRI and histology is performed via MI and COFEMI.
2. Subsequent TPS warping is applied following affine alignment to achieve the final image alignments (MI+TPS and COFEMI+TPS) using prominent anatomical landmarks as control points.
3. The ground truth for spatial extent of cancer is directly mapped onto the CAD-targeted modality (MRI).
4. Manual delineations of cancer on MRI are obtained from radiologists using visual cues from histology.

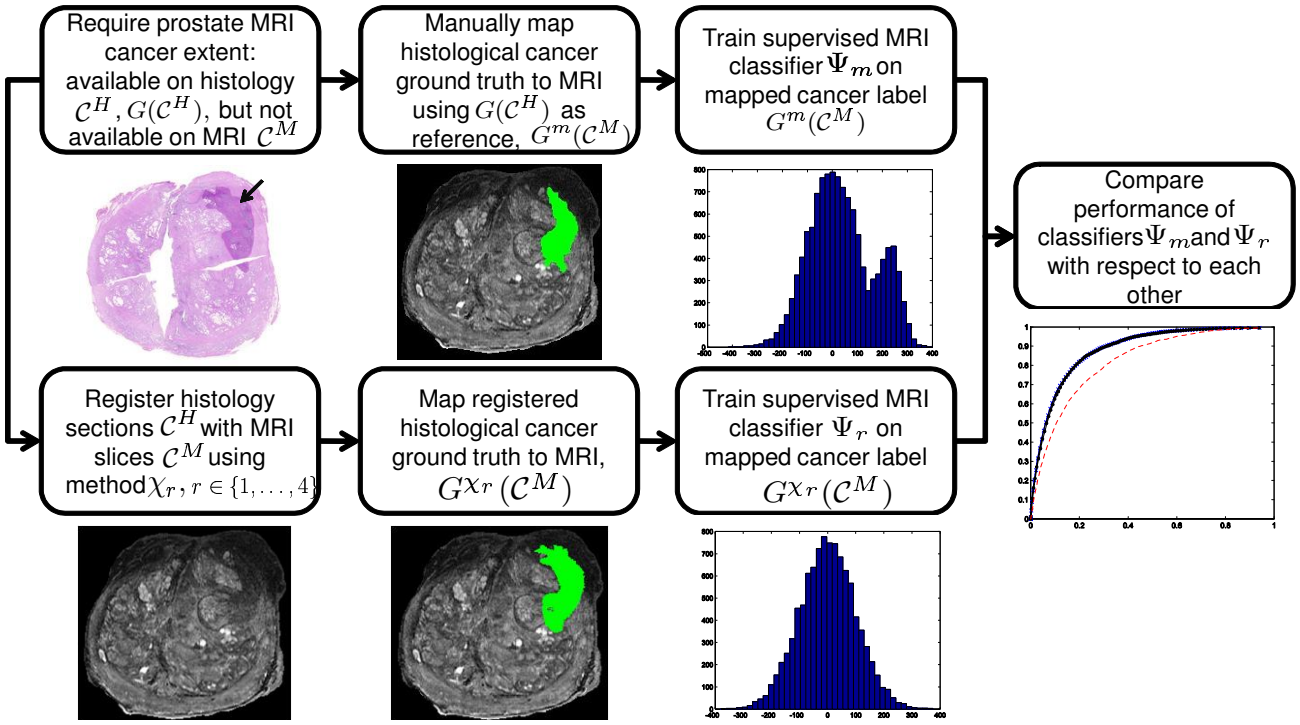


Figure 2. Outline of registration-based supervised classifier training system applied to CAD for prostate MRI with histological ground truth for cancer (top left). Labeling of MRI training data is achieved by registration of histology with MRI, followed by directly mapping histological ground truth to MRI, which are used to train the supervised classifier which drives the CAD system (bottom row). Training labels are also established by expert manual delineation, upon which another classifier is trained (top row). The classifiers are then evaluated with respect to each other (right).

5. The pixels automatically labeled as cancer and those delineated by experts are used as training instances in separate supervised classifiers.
6. Accuracies of the classifiers trained using the 4 registration methods and the expert delineation are compared in terms of ROC curves using a 5-fold randomized cross validation procedure.

In this manner, we demonstrate that in spite of the challenges in registration of prostate MRI and histology, the automated CaP annotations thus obtained are significantly more accurate compared to expert delineation of CaP on MRI obtained with aid from corresponding stained histological sections and prior domain knowledge. We further demonstrate that despite the non-linear nature of the deformations even simple rigid registration methods are more accurate compared to expert annotation, yielding better quality classifier training data.

3. OVERVIEW OF REGISTRATION METHODS

3.1 Notation

We define a 2D image by the notation $\mathcal{C} = (C, f)$ where C is a coordinate grid, called the image domain, and f is a function that assigns to every spatial location $\mathbf{c} \in C$ an intensity value $f(\mathbf{c})$ in 2D space $\mathbf{c} = (x, y)$. Under this convention, the 2D slices of the 3D *ex vivo* MRI dataset are defined by $\mathcal{C}^M = (C, f^M)$, while the corresponding histology slices are defined by $\mathcal{C}^H = (C, f^H)$. The notations and symbols used in the rest of this paper are listed in Table 1.

3.2 Manual Mapping of Cancer from Histology to MRI

Histological sections were microscopically examined to determine cancer extent on $\mathcal{C}^H$ denoted by the set of pixels $G(\mathcal{C}^H)$. Manual expert segmentations of cancer on *ex vivo* MRI, denoted by the set of pixels $G^m(\mathcal{C}^M)$, are obtained by visually mapping $G(\mathcal{C}^H)$ onto $\mathcal{C}^M$. Note that the manual delineation of cancer $G^m(\mathcal{C}^M)$ on $\mathcal{C}^M$ was guided by $G(\mathcal{C}^H)$ on $\mathcal{C}^H$ along with visual cues from the anatomical features of the prostate on both $\mathcal{C}^H$

Table 1. List of notation and symbols used in this paper.

Symbol	Description	Symbol	Description
$\mathcal{C}^M$	2D MRI image	$\mathcal{C}^H$	2D Histology image
C	Set of 2D pixel coordinates $\mathbf{c}$	$f(\mathbf{c})$	Intensity value at pixel $\mathbf{c}$
$G(\mathcal{C}^H)$	CaP pixels from H&E staining on $\mathcal{C}^H$	$G^m(\mathcal{C}^M)$	CaP pixels found manually on $\mathcal{C}^M$
χ_r	Registration method where $r \in \{1,2,3,4\}$	$G^{\chi_r}(\mathcal{C}^M)$	$G(\mathcal{C}^H)$ on $\mathcal{C}^M$ by method χ_r
$\mathcal{C}^{\Phi_\beta}$	MRI feature image for operator Φ_β	Φ_β	Feature operator where $\beta \in \{1,\ldots,n\}$
$f^{\Phi_\beta}(\mathbf{c})$	Feature value at $\mathbf{c}$ for Φ_β	$\mathbf{f}(\mathbf{c})$	Feature vector $[f^{\Phi_\beta}(\mathbf{c})\|\beta \in \{1,\ldots,n\}]$ at $\mathbf{c}$

and $\mathcal{C}^M$. Since there were roughly twice the number of MRI slices as there were histological sections for each patient study, the MRI sections corresponding to the histological sections were manually identified by an expert by identifying anatomical features visible on both $\mathcal{C}^H$ and $\mathcal{C}^M$. A total of 26 MRI-histological sections were considered in this study.

3.3 Overview of Registration Methods

Most robust rigid/affine multimodal registration techniques utilize similarity measures to bring images into alignment, the assumption being that when the images are in alignment, their similarity is maximized. The optimal alignment of the two images $\mathcal{A}$ and $\mathcal{B}$ obtained by the transformation T is given by

$$T_{max} = \underset{T}{\operatorname{argmax}}\, \psi(T\mathcal{A}, \mathcal{B}), \tag{1}$$

where ψ is some voxel-based similarity measure and $T\mathcal{A}$ represents the image $\mathcal{A}$ subject to transformation T. The designation of $\mathcal{A}$ and $\mathcal{B}$ as moving (target) and stationary (template) images is user dependent.For the case of $\mathcal{A}$ as the target of deformations, we further denote $\hat{\mathcal{A}}$ as the result of the initial affine registration step and $\tilde{\mathcal{A}}$ as the result of the subsequent elastic registration step. The four effective affine and elastic registration techniques (denoted $\chi_r \in$ {MI, MI+TPS, COFEMI, COFEMI+TPS}) evaluated in this study are described in the following sections.

3.3.1 Maximization of Mutual Information (MI)

The most common similarity measures used for voxel-wise image registration are the sums of squared distance (SSD), correlation coefficient (CC), mutual information (MI), related information theoretic quantities such as entropy correlation coefficient (ECC), normalized mutual information (NMI).[13] Since no linear relationship between voxel intensities in $\mathcal{A}$ and $\mathcal{B}$ may exist with multimodal imagery, mutual information and other statistical measures of image similarity or interdependence have been widely utilized for multimodal image registration tasks. Registration by maximization of MI can be achieved by calculating the MI of two images in terms of Shannon entropy,[13] a measure of information content of a random variable. Equation 2 is a formulation of MI in terms of the marginal and joint entropies wherein the MI of a pair of images or random variables, $I_2(\mathcal{A},\mathcal{B})$, is maximized by minimizing joint entropy $S(\mathcal{AB})$ and maintaining the marginal entropies $S(\mathcal{A})$ and $S(\mathcal{B})$. Where ensemble $\mathcal{AB}$ represents simultaneous knowledge of both images, MI is given by,

$$I_2(\mathcal{A},\mathcal{B}) = S(\mathcal{A}) + S(\mathcal{B}) - S(\mathcal{AB}), \tag{2}$$

where $I_2(\mathcal{A},\mathcal{B})$ describes the interdependence of multiple variables or gray levels of a set of images.[13] Thus, when $I_2(\mathcal{A},\mathcal{B})$ increases, the uncertainty about $\mathcal{A}$ given $\mathcal{B}$ decreases. As per Equation 1, it is assumed that the global MI maximum will occur at the point of precise registration.

3.3.2 Combined Feature Ensemble Mutual Information

Both MI and NMI can be ill-defined and highly non-convex measures for dissimilar modalities.[13–15] Hence, in [12] we presented the combined feature ensemble mutual information (COFEMI) registration technique, which incorporates additional information in the form of image features which are complementary to image intensity in order to increase the robustness of the similarity measure to modality differences and image artifacts, and was demonstrated to be superior compared to MI in terms of robustness to images with artifacts and dissimilar

modalities. The COFEMI technique utilizes the notion of combined mutual information (CMI) to incorporate additional information in the form of image features. The CMI that two semi-independent images, $\mathcal{B}$ and $\mathcal{B}'$, contain about a third image, $\mathcal{A}$, is defined as $I_2(\mathcal{A}, \mathcal{B}\mathcal{B}') = S(\mathcal{A}) + S(\mathcal{B}\mathcal{B}') - S(\mathcal{A}\mathcal{B}\mathcal{B}')$.[16] This formulation allows the incorporation of only the unique (non-redundant) information provided by an additional image, $\mathcal{B}'$, about $\mathcal{A}$. Hence, the generalized form of CMI for n additional images is,

$$I_2(\mathcal{A}, \mathcal{B}\mathcal{B}'_1 \cdots \mathcal{B}'_n) = S(\mathcal{A}) + S(\mathcal{B}\mathcal{B}'_1 \cdots \mathcal{B}'_n) - S(\mathcal{A}\mathcal{B}\mathcal{B}'_1 \cdots \mathcal{B}'_n). \quad (3)$$

It can be seen that CMI incorporates only the unique information of additional images toward describing $\mathcal{A}$, thus enhancing but not overweighting the similarity measure with redundant information. Therefore, it will always be the case that $I_2(\mathcal{A}, \mathcal{B}\mathcal{B}'_1 \cdots \mathcal{B}'_n) \leq S(\mathcal{A}) = I_2(\mathcal{A}, \mathcal{A})$. The intuition behind using CMI is that one or more of the feature images $\mathcal{B}'_1, \mathcal{B}'_2, ..., \mathcal{B}'_n$ derived from the intensity image $\mathcal{B}$ will be relatively more immune to image intensity artifacts such as bias field inhomogeneity and intensity non-linearity compared to $\mathcal{B}$. In addition one or more of $\mathcal{B}'_1, \mathcal{B}'_2, ..., \mathcal{B}'_n$ will potentially provide additional structural description of $\mathcal{A}$ not readily discernible from intensity image $\mathcal{B}$. By choosing appropriate statistical features calculated over windows of $\mathcal{B}$, one may introduce information pertaining to the spatial arrangement of intensity values into the similarity measure. The feature extraction and selection algorithms described in [12] were used to generate a feature ensemble $\pi_k \subset \{\mathcal{B}'_1, \mathcal{B}'_2, ..., \mathcal{B}'_n\}$ of length k used to drive the registration procedure by maximization of $I_2(\mathcal{A}, \mathcal{B}\,\pi_k)$. Details of feature extraction are given in Section 4.1.

3.3.3 Affine Registration via MI and COFEMI

Images $\mathcal{A}$ and $\mathcal{B}$ are initially aligned using an affine transformation with a total of five parameters for rotation (θ), translation (dx, dy), and scaling (σ_x, σ_y). The affine transformation matrix ($\mathbf{M}$) is constructed by combination of three coordinate transformation matrices corresponding to rotation in the plane of the image ($\mathbf{R}$), translation in the X and Y directions ($\mathbf{T_{xy}}$), and horizontal and vertical scaling ($\mathbf{S_{xy}}$). The individual affine transformation matrices are composed in terms of the five parameters ($\theta, dx, dy, \sigma_x, \sigma_y$) by common geometric coordinate transformations equations. Homogeneous transformation matrices and coordinate representations were utilized to apply translations in the same matrix operation as scaling and rotation. Thus for a homogeneous coordinate $\mathbf{c_h} = (x, y, 1)^T$, the transformed coordinate is determined by $\mathbf{c'_h} = \mathbf{M}\ \mathbf{c_h}$. Deformations are applied in the order of rotation, scaling, and then translation by the product $\mathbf{M} = \mathbf{T_{xy}}\ \mathbf{S_{xy}}\ \mathbf{R}$. Both COFEMI and MI registration techniques (χ_1 and χ_2) utilize NN interpolation to avoid artifacts associated with linear and partial volume interpolation. MI-based registration is achieved by optimizing (via a Nedler-Mead simplex algorithm tuned for differences in parameter scale) the affine transformation $\mathbf{M}$ by maximization of MI of $\mathcal{A}$ with $\mathcal{B}$ to obtain the optimal tranformation $\mathbf{M}^{\chi_1}$. The COFEMI procedure operated by maximization of CMI of $\mathcal{A}$ and ensemble $\mathcal{B}\,\pi_k$, resulting in the optimal rigid transformation $\mathbf{M}^{\chi_2}$ to best align $\mathcal{B}$ with $\mathcal{A}$ by maximizing $I_2(\hat{\mathcal{A}}, \mathcal{B}\,\pi_k)$. The common coordinate frame C is maintained.

3.3.4 Elastic Registration via MI, COFEMI, and TPS

Thin plate splines (TPS)[17] were used to model the non-linear portion of deformation required to bring $\mathcal{A}$ and $\mathcal{B}$ into alignment. That is, having obtained $\hat{\mathcal{A}}$ via affine registration, a TPS defined deformation field was used to generate the final alignment result $\tilde{\mathcal{A}}$ from $\hat{\mathcal{A}}$. Using the approach outlined by Bookstein,[17] the equation describing the surface of a thin material plate is used as an interpolant to define in-plane deformations based on the translation of arbitrary control points. Thus, to define a warping of image $\hat{\mathcal{A}}$ onto image $\mathcal{B}$, two sets of control points $\mathbf{c}$ and $\mathbf{d}$ representing the same anatomical features on images $\mathcal{A}$ and $\mathcal{B}$, respectively, are selected by manually identifying salient anatomical landmarks. The landmarks are structural features that are clearly visible on both modalities, such as large duct centers, the urethra, features along the outer boundary of the prostate and along the transition between peripheral and central zones. For any slice pair, $\{\mathbf{c}_1, \ldots, \mathbf{c}_p\}, \{\mathbf{d}_1, \ldots, \mathbf{d}_p\} \subset C$ where the number of control point pairs rage from $7 \leq p \leq 12$. Anatomical control points are chosen on the common undeformed images such that precisely the same anatomical features are used to define the elastic deformations that follow the different affine alignments obtained by MI and COFEMI. The principal warp determined from the solution of the thin plate spline subject to the displacement of points $\mathbf{c}$ and $\mathbf{d}$ is applied to obtain the overall affine and elastic deformation fields $\mathbf{M}^{\chi_3}$ and $\mathbf{M}^{\chi_4}$. Similarly, $\tilde{\mathcal{B}}$ may be obtained from $\hat{\mathcal{B}}$ if $\mathcal{A}$ is desired to remain stationary. We refer to elastic registration following MI- and COFEMI-based affine registration as MI+TPS and COFEMI+TPS.

4. SUPERVISED COMPUTER-AIDED CLASSIFICATION

4.1 Feature Extraction

We compute a total of $n = 311$ unique feature images from each MRI intensity image $\mathcal{C}^M$, the CAD-targeted modality, which corresponds to image $\mathcal{B}$ in the generalized notation used throughout Section 3.3 above. These feature representations were chosen for their class-discriminative abilities as previously demonstrated in prostate MRI CAD.[1] These features were also chosen to compose the ensemble used in the COFEMI registration scheme described in Section 3.3.2 since several of these feature representations were found to be less susceptible to image artifacts such as bias field inhomogeneity.[12] Thus, for each $\mathcal{B}$, feature images $\mathcal{B}^{\Phi_\beta} = (C, f^{\Phi_\beta})$ are calculated by applying feature operators Φ_β, $\beta \in \{1, \ldots, n\}$ within a local neighborhood associated with every $\mathbf{c} \in C$. Thus $f^{\Phi_\beta}(\mathbf{c})$ represents the feature value associated with feature operator Φ_β at pixel $\mathbf{c}$ in $\mathcal{B}$, and $\mathbf{f} = [f^{\Phi_\beta}(\mathbf{c}) | \beta \in \{1, \ldots, n\}]$ is the feature vector of length n for pixel $\mathbf{c}$.

4.1.1 Gradient Features

Gradient features are calculated using steerable and non-steerable linear gradient operators. Eleven non-steerable gradient features were obtained using Sobel, Kirsch and derivative operations. Gabor gradient operators[18] comprise the steerable class of gradient calculations. The magnitude of the Gabor filter response was defined for every pixel $\mathbf{c} = (x, y)$ by $f^{Gab}(\mathbf{c}) = e^{(x'+y')/2\pi} \cos(2\pi\lambda x')$ where λ is the spatial frequency (scale) of the sinusoid. The filter was applied on a $\kappa \times \kappa$ square neighborhood $\mathcal{N}_{\mathbf{c},\kappa}$ associated with every pixel $\mathbf{c} \in C$ and centered on $\mathbf{c}$. The orientation of the filter is affected by φ through the coordinate transformations: $x' = x\cos\varphi + y\sin\varphi$ and $y' = x\cos\varphi - y\sin\varphi$. Gabor gradient features were calculated for $\kappa \in \{5, 8, 15\}$, $\varphi \in \{0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{8}, \frac{3\pi}{4}, \frac{7\pi}{8}\}$ and $\lambda \in \{\frac{\pi}{2\sqrt{2}}, \frac{\pi}{4}, \frac{\pi}{4\sqrt{2}}, \frac{\pi}{8}, \frac{\pi}{8\sqrt{2}}, \frac{\pi}{16}\}$.

4.1.2 Statistical Features

Four first order statistical features, including mean, median, standard deviation, and range, and were calculated as described in [1]. To calculate the second order statistical feature images, we define any pixel $\mathbf{d} \in \mathcal{N}_{\mathbf{c},\kappa}$ as a κ neighbor of $\mathbf{c}$. We now compute a $M \times M$ co-occurrence matrix $P_{\mathbf{c},\kappa}$ associated with $\mathcal{N}_{\mathbf{c},\kappa}$, where M is the maximum gray scale intensity over all $\mathcal{B}$. The value at any location $[u, v]$, where $u, v \in \{1, \ldots, M\}$, in the matrix $P_{\mathbf{c},\kappa}[u, v]$ represents the frequency with which two distinct pixels $\mathbf{c}, \mathbf{d} \in \mathcal{N}_{\mathbf{c},\kappa}$ with associated image intensities $f(\mathbf{c}) = u$, $f(\mathbf{d}) = v$ are adjacent (within the same 8-neighborhood within $\mathcal{N}_{\mathbf{c},\kappa}$). A total of 16 Haralick features including energy, entropy, inertia, correlation, inverse difference moment, two information correlation measures, sum average, sum variance, sum entropy, different average, difference variance, difference entropy, shade, prominence and variance were extracted at every pixel $\mathbf{c} \in C$, from $P_{\mathbf{c},\kappa}$, for $\kappa \in \{5, 8, 15\}$, and $M \in \{64, 128, 256\}$.

4.2 Bayesian Classifier and Independent Component Analysis

The naïve form of Bayes' rule defines the likelihood of observing some class ω_1 given the multivariate feature vector $\mathbf{f}(\mathbf{c})$ of dimensionality n and is determined by,

$$P(\omega_1|\mathbf{f}(\mathbf{c})) = \frac{P(\omega_1)p(\mathbf{f}(\mathbf{c})|\omega_1)}{P(\omega_1)p(\mathbf{f}(\mathbf{c})|\omega_1) + P(\omega_0)p(\mathbf{f}(\mathbf{c})|\omega_0)}. \quad (4)$$

where the prior probabilities of occurrence of the two classes are $P(\omega_1)$ and $P(\omega_0)$, and $p(\mathbf{f}(\mathbf{c})|\omega_1)$ and $p(\mathbf{f}(\mathbf{c})|\omega_0)$ are the *a priori* class conditional distributions of $\mathbf{f}(\mathbf{c})$. Estimation of $p(\mathbf{f}(\mathbf{c})|\omega_1)$ and $p(\mathbf{f}(\mathbf{c})|\omega_0)$ is difficult or impossible when the distributions are of a high dimensionality. However, if the conditional probabilities of each dimension of $\mathbf{f}(\mathbf{c})$ can be made statistically independent by some linear transformation W (i.e. $\boldsymbol{\phi} = W\mathbf{f}$) the distribution of the new random variable $\boldsymbol{\phi}$ of dimensionality h can be written $p(\boldsymbol{\phi}) = p(\phi_1, \ldots, \phi_k) = \prod_i^h p(\phi_i)$. Hence, independent component analysis (ICA) is used to reduce the feature vector $\mathbf{f}$ of size $n = 311$ to a vector $\boldsymbol{\phi} = \{\phi_1, \ldots, \phi_h\}$ of size $h \ll n$ with statistically independent distributions. That is, for the entire set of pixels C of image $\mathcal{B}$, ICA determines a $h \times n$ unmixing matrix W such that the distributions of the rows of $\boldsymbol{\phi}$ are maximally independent. The h independent components $\phi_i(\mathbf{c})$ of pixels $\mathbf{c} \in G^{\chi_r}$ are then used to generate h *a priori* distributions $p(\phi_i(\mathbf{c})|\omega_1)$ for the target class ω_1, while pixels $\mathbf{c} \in \{C - G^{\chi_r}\}$ are used to generate

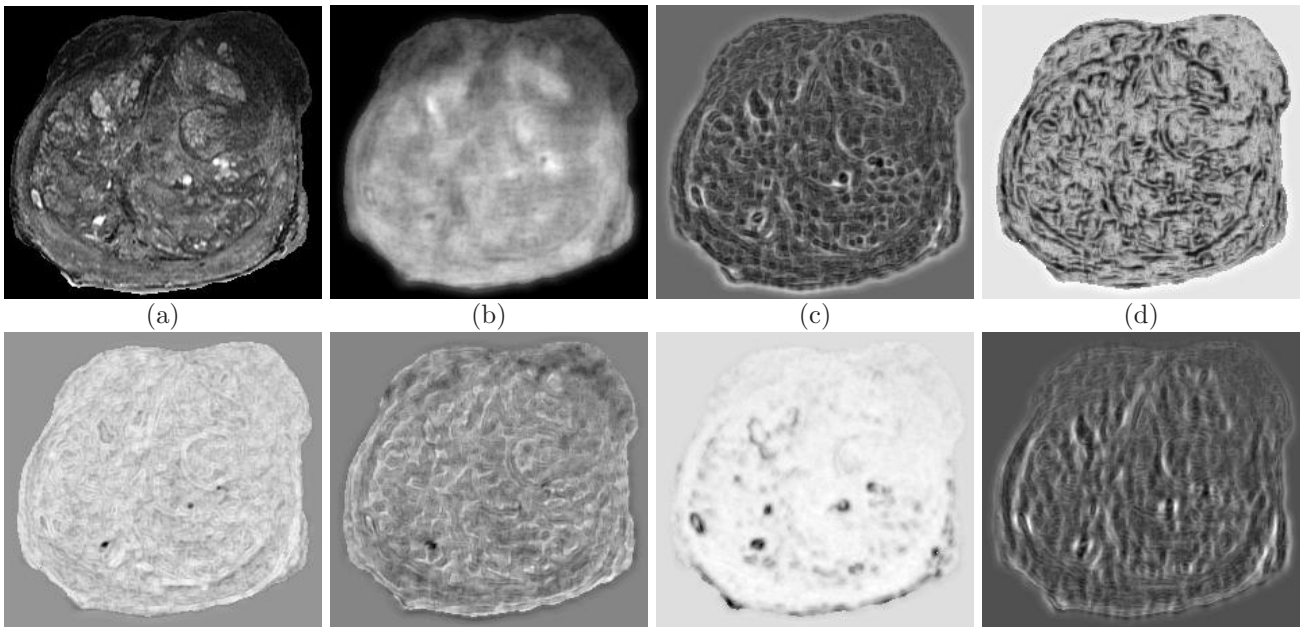


Figure 3. MRI intensity image (a) and corresponding independent component feature images obtained via ICA (b)-(h) of feature images (not shown). Maximal statistical independence between independent components is maximized to facilitate naïve combination in Bayes' rule.

distributions for class ω_0. Hence, Bayes' rule can then be used to obtain the *a posteriori* probability distribution $P(\omega_1|\boldsymbol{\phi}(\mathbf{c}))$ of observing ω_1 for the linearly independent feature vector $\boldsymbol{\phi}(\mathbf{c})$ at each pixel $\mathbf{c} \in C$ by

$$P(\omega_1|\boldsymbol{\phi}(c)) = \frac{P(\omega_1)\prod_i^h p(\phi_i(\mathbf{c})|\omega_1)}{P(\omega_1)\prod_i^h p(\phi_i(\mathbf{c})|\omega_1) + P(\omega_0)\prod_i^h p(\phi_i(\mathbf{c})|\omega_0)}. \tag{5}$$

The probability images $\mathcal{L} = (C, P(\omega_1|\boldsymbol{\phi}(\mathbf{c})))$, represent the joint *a posteriori* likelihood of $\mathbf{c} \in C$ belonging to class ω_1 obtained for each training image $\mathcal{B}$, which correspond to MRI slices ($\mathcal{C}^M$) in this study. Importantly, since the class of features represented by the individual feature operators Φ_β has been demonstrated to be discriminative in previous studies,[1] the independent component distributions $P(\omega_1|\phi_i(\mathbf{c}))$ will contain discriminative potential.

The Jade ICA implementation[19] was used to employ a measure of independence in approximation to MI, a natural measure of statistical dependence of multiple distributions related to measures of nongaussianity that are commonly used in ICA algorithms. In this view of independence, the maximum number of independent component images that can be obtained from the feature set with a minimal degree of dependence between the components is determined via calculation of pairwise MI between the resulting independent component images as described in [20]. Estimation of further "independent" components beyond the maximum results in a sharp increase in the magnitude of MI between multiple components, indicating that additional independent sources are not identifiable in the feature set. In this manner, the value of h is determined. Figure 3(b)-(h) shows $h = 7$ independent component images derived from feature images of $\mathcal{C}^M$ in Figure 3(a), each of which were tested for independence by paired MI calculations and inspection of the joint distributions.

5. APPLICATION TO PROSTATE CAD

5.1 Supervised Classifier Training

The registration-based supervised classifier training technique described above was used to train a CAD system for detection of cancer in MRI images of the prostate ($\mathcal{C}^M$). In the registration procedure we treat $\mathcal{C}^M$ as the stationary template image ($\mathcal{A}$) and $\mathcal{C}^H$ as the target image ($\mathcal{B}$). The four effective affine and elastic registration

techniques ($\chi_r \in$ {MI, MI+TPS, COFEMI, COFEMI+TPS}) described in Section 3.3 were each used to establish corresponding region labels $G^{\chi_1}(\mathcal{C}^M)$, $G^{\chi_2}(\mathcal{C}^M)$, $G^{\chi_3}(\mathcal{C}^M)$ and $G^{\chi_4}(\mathcal{C}^M)$ on MRI. The region labels were each used to establish the membership of each pixel in $\mathcal{C}^M$ to either the cancer or non-cancer class ($\mathbf{c} \hookrightarrow \omega_c$ or $\mathbf{c} \hookrightarrow \omega_{nc}$), and four supervised Bayesian classifiers Ψ_1, Ψ_2, Ψ_3 and Ψ_4 were trained as described in Section 4.2. The manual delineation of cancer $G^m(\mathcal{C}^M)$ was used to train a fifth Bayesian classifier Ψ_5.

5.2 Evaluation of Supervised CAD Classifier

Since no definitive ground truth is available on $\mathcal{C}^M$, testing and evaluation of classifier accuracy was done using seven ground truth surrogates, which correspond to the five individual cancer labels ($G^{\chi_1}(\mathcal{C}^M)$, $G^{\chi_2}(\mathcal{C}^M)$, $G^{\chi_3}(\mathcal{C}^M)$ and $G^{\chi_4}(\mathcal{C}^M)$ and $G^m(\mathcal{C}^M)$), the union ($\bigcup G^\Omega$ where $\Omega \in \{\chi_1, \chi_2, \chi_3, \chi_4, m\}$) and the intersection ($\bigcap G^\Omega$) of the labels. These 7 ground truth estimates are represented as $\mathbf{\Gamma}_{1-7}$. For each classifier Ψ_{1-5} and ground truth estimate $\mathbf{\Gamma}_{1-7}$, totaling 35 combinations, receiver operating characteristic (ROC) curves are generated, and the area under the ROC curves (AUC) and positive predictive value (PPV) are calculated. The vertical axis of the ROC curve is the true positive rate (TPR) or sensitivity and the horizontal axis is the false positive rate (FPR) or 1-specificity. Each point on the curve corresponds to the sensitivity and specificity of detection of classifier $\Psi_\alpha, \alpha \in \{1, \ldots, 5\}$, at threshold ρ. If $P(\omega_c|\boldsymbol{\phi}(\mathbf{c})) \geq \rho$ then $\mathbf{c}$ is classified as ω_c else ω_{nc} to obtain a binary prediction $G^{CAD}_{\rho,\Psi}$ for threshold ρ and classifier Ψ. For each $G^{CAD}_{\rho,\Psi}$ and $\mathbf{\Gamma}_{1-7}$, sensitivity (SN) and specificity (SP) are calculated as

$$SN_{\rho,\Psi,\Gamma} = \frac{\left|G^{CAD}_{\rho,\Psi} \cap \mathbf{\Gamma}\right|}{\left|G^{CAD}_{\rho,\Psi} \cap \mathbf{\Gamma}\right| + \left|\mathbf{\Gamma} - G^{CAD}_{\rho,\Psi} \cap \mathbf{\Gamma}\right|} \text{ and } SP_{\rho,\Psi,\Gamma} = \frac{\left|(C - G^{CAD}_{\rho,\Psi}) \cap (C - \mathbf{\Gamma})\right|}{\left|(C - G^{CAD}_{\rho,\Psi}) \cap (C - \mathbf{\Gamma})\right| + \left|G^{CAD}_{\rho,\Psi} - \mathbf{\Gamma}\right|} \tag{6}$$

where $|.|$ indicates the cardinality of a set of pixels. A 5-fold randomized cross-validation procedure is used to compare the accuracy of each Ψ_{1-5} with respect to each $\mathbf{\Gamma}_{1-7}$, whereby from the 26 image pairs with $G(\mathcal{C}^H)$ mapped to $\mathcal{C}^M$ by each method, 21 MR images are randomly chosen from which a classifier is trained. The remaining images are used as a test set upon which classifier accuracy is evaluated. Randomized cross-validation is repeated 50 times for different training and testing slices $\mathcal{C}^M$.

Average ROC curves for each classifier were generated for display purposes by fitting a smooth polynomial and averaging the ROCs over each test image and $\mathbf{\Gamma}_{1-7}$. To avoid introducing bias toward classifiers derived from registration-established training data, which would occur if curves from all four $\mathbf{\Gamma}_{1-4}$ are averaged, each average curve was generated using a subset of $\mathbf{\Gamma}_{1-7}$: the expert determined ground truth ($\mathbf{\Gamma}_5$), the intersection and union ground truths ($\mathbf{\Gamma}_{6-7}$) and one automatically mapped ground truth estimate ($\mathbf{\Gamma}_{1-4}$ corresponding to the Ψ being evaluated). Values of AUC and PPV for each classifier pair were compared in a similar manner.

6. RESULTS

6.1 Qualitative Evaluation of Registration Schemes

For each of the 26 histological and MRI images, alignment of corresponding $\mathcal{C}^M$ and $\mathcal{C}^H$ slices and establishment of training instances in $\mathcal{C}^M$ was performed using each of the four registration techniques. Figure 4 shows the results of the registration procedures for a single corresponding pair of images $\mathcal{C}^H$ (Fig. 4(a)) and $\mathcal{C}^M$ (Fig. 4(e)). The manual expert delineation of cancer on MRI ($G^m(\mathcal{C}^M)$) is shown in green in Fig. 4(b). Histological ground truth $G(\mathcal{C}^H)$, indicated by the arrow in Fig. 4(a), is shown mapped to $G^{\chi_1}(\mathcal{C}^M)$, $G^{\chi_4}(\mathcal{C}^M)$ and $G^{\chi_3}(\mathcal{C}^M)$ in Figs. 4(c),(d) and (f). Comparing $G^{\chi_1}(\mathcal{C}^M)$ and $G^{\chi_3}(\mathcal{C}^M)$ as shown in Figs. 4(c) and (f) indicates similar results with fewer obvious mislabeled instances in $G^{\chi_3}(\mathcal{C}^M)$. However, the difference in accuracy of the registration methods is best illustrated as an overlay of $\mathcal{C}^M$ with the whole transformed image $\mathcal{C}^H$ as in Figs. 4(g) and (h) for the methods MI and COFEMI+TPS. Notice the considerable region of non-overlapping tissue in Figure 4(g) compared with the nearly ideal alignment in Figure 4(h). It is clear that a non-linear elastic registration procedure is superior in this task. Over all slice pairs, COFEMI was found to produce qualitatively more accurate alignments compared with MI, supporting previous results.[12] Further, the associated elastic methods MI+TPS and COFEMI+TPS produced the most qualitatively accurate and realistic alignments overall.

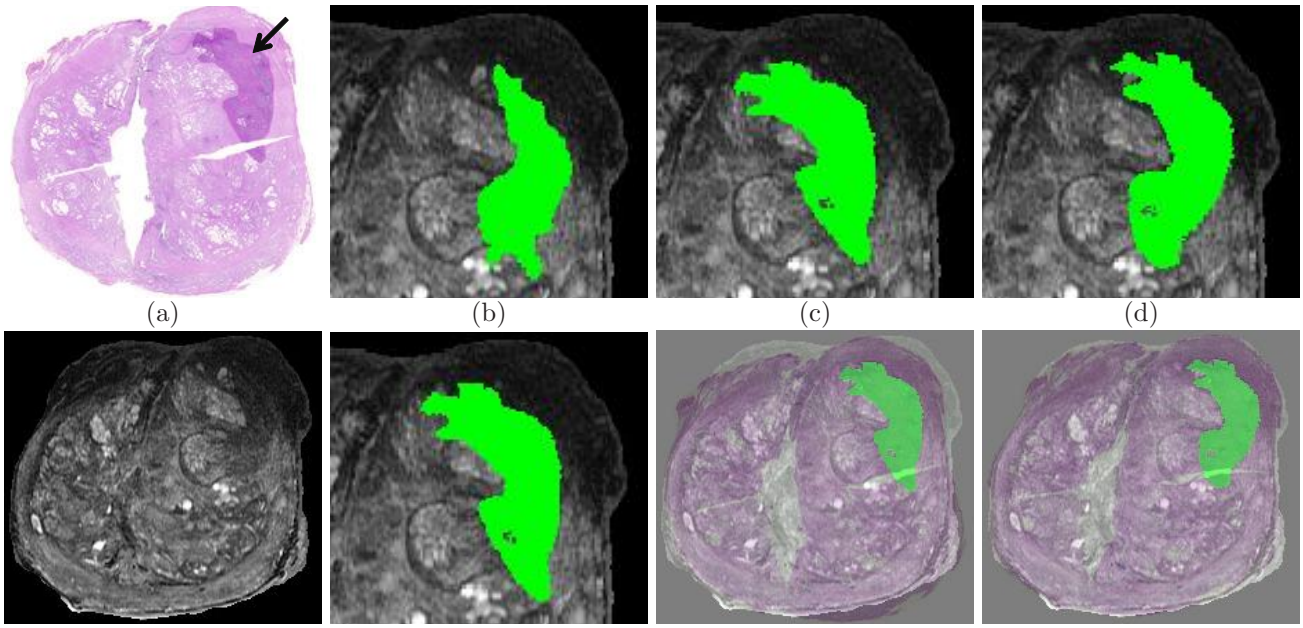


Figure 4. (a) Histology with ground truth indicated by the black arrow is registered to (e) MRI to transfer histological cancer map to MRI. Resulting mappings obtained via MI (c), COFEMI+TPS (d) and COFEMI (f). Overlays of MRI with histology registered via MI and COFEMI+TPS shown in (g) and (h), respectively, indicate the more accurate alignment with elastic vs. affine deformations. (b) Manually determined labels on MRI are shown for comparison.

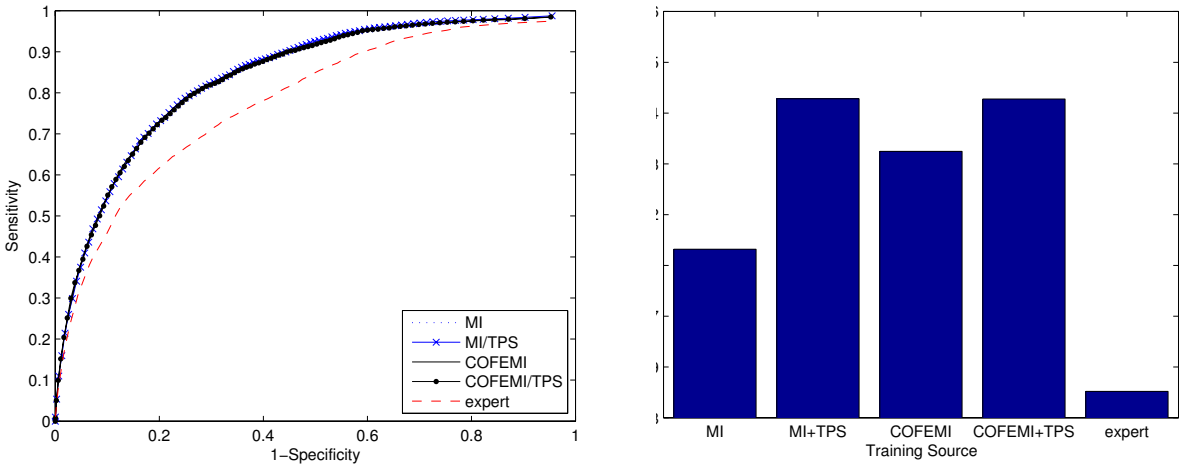
Figure 4h http://dx.doi.org/10.1117/12.770703.1 Figure 4 a and d: http://dx.doi.org/10.1117/12.770703.2

6.2 Quantitative Results

From the registration results, the supervised classifiers Ψ_{1-5} were generated using $h = 7$ independent component images obtained from the full set of feature images. Using the randomized cross-validation procedure described above, Ψ_{1-5} were compared for the ground truth estimates $\mathbf{\Gamma}_{1-7}$ in terms of PPV and AUC. Figure 5(a) shows average ROC curves for each classifier Ψ_{1-5} derived from the different training data labels ($G^{\chi_1}(\mathcal{C}^M)$, $G^{\chi_2}(\mathcal{C}^M)$, $G^{\chi_3}(\mathcal{C}^M)$, $G^{\chi_4}(\mathcal{C}^M)$ and $G^{m}(\mathcal{C}^M)$). The dashed line in Figure 5 represents the ROC curve for classifier Ψ_5 averaged over different CaP ground truth estimates as described in Section 5.2. Clearly, overall accuracy of the classifiers Ψ_{1-4} trained from registration-derived training labels is greater than the accuracy of Ψ_5, which is trained using expert labels. Values for AUC and PPV were then calculated from the original ROC curves. Differences between classifiers Ψ_{1-5} are compared in terms of PPV values as illustrated in Figure 5(b), which shows the average PPV for each Ψ_{1-5}. The difference in PPV between Ψ_5 (expert) and each Ψ_{2-4} (MI+TPS, COFEMI, COFEMI+TPS) were statistically significant ($p < 0.05$) for each $\mathbf{\Gamma}_{1-4,6-7}$, while for $\mathbf{\Gamma}_5$ the values for Ψ_5 were not significantly greater than any Ψ_{1-4}. The AUC of each Ψ_{1-4} (MI, MI+TPS, COFEMI, COFEMI+TPS) was significantly ($p < 0.05$) greater than the AUC of Ψ_5 (expert) for the same $\mathbf{\Gamma}$s as for PPV, while Ψ_5 again displayed no significant advantage for any $\mathbf{\Gamma}$.

7. CONCLUDING REMARKS

In this paper we demonstrated the need for use of multi-modal registration methods for establishing accurate ground truth for classifier training. We demonstrate that even simple MI-based schemes are more accurate in mapping spatial extent of cancer from one modality to another compared to expert labeling in spite of inter-modal differences and intensity artifacts which hinder automated registration techniques. We quantitatively demonstrated with the use of 4 different registration methods (MI, COFEMI, MI+TPS, COFEMI+TPS) to align 26 pairs of *ex vivo* MRI with histology images of the prostate upon which cancer ground truth is available and provide improved registration accuracy compared to expert delineations. The 5 labels for CaP were used to train 5 classifiers and these were evaluated against 7 different ground truth surrogates. A quantitative performance improvement in the supervised classifier component of a CAD system was observed using the cancer labels



(a) (b)

Figure 5. Accuracy of classifiers Ψ_{1-5} represented by the average ROC curves in (a). Quantitative comparison of Ψ_{1-5} in terms of AUC (not shown) indicates a statistically significant improvement from training labels established by registration vs. expert delineation. Comparison in terms of (b) PPV suggests an improvement in labeling using elastic registration.

obtained by registration versus labels manually determined by an trained expert pathologist with a histological reference.

This method for training of a supervised classifier by multimodal registration is applicable to any scenario in which multimodal data with ground truth is available. By providing accurate and repeatable training labels, registration overcomes several of the problems with manual delineation of training labels including human error, observer bias and variability. Further, the method of comparing classifier performance where the classifiers have been trained on the registration results may be used as a means of comparing registration performance. Some steps in this direction were undertaken in this paper.

ACKNOWLEDGMENTS

This work was funded by: Department of Defense Prostate Cancer Research Program (W81XWH-08-1-0072), The Coulter Foundation (WHCF4-29349, WHCF 4-29368), Busch Biomedical Award, Cancer Institute of New Jersey, New Jersey Commission on Cancer Research, National Cancer Institute (R21-CA127186-01, R03-CA128081-01).

REFERENCES

1. A. Madabhushi, M. D. Feldman, *et al.*, "Automated detection of prostatic adenocarcinoma from high-resolution *ex vivo* MRI," *IEEE Trans. Med. Imag.* **24**, pp. 1611–1625, December 2005.
2. A. Madabhushi, J. Shi, *et al.*, "Graph embedding to improve supervised classification and novel class detection: application to prostate cancer," in *MICCAI 2005*, J. S. Duncan and G. Gerig, eds., pp. 729–737, 2005.
3. C. E. Brodley and M. A. Friedl, "Identifying mislabeled training data," *J. Artif. Intell. Res.* **11**, pp. 131–167, 1999.
4. F. Muhlenbach, S. Lallich, and D. A. Zighed, "Identifying and handling mislabelled instances," *J. Intell. Inf. Syst.* **22**(1), pp. 89–109, 2004.
5. L. S. Taylor, B. C. Porter, *et al.*, "Three-dimensional registration of prostate images from histology and ultrasound.," *Ultrasound Med Biol* **30**, pp. 161–168, Feb 2004.
6. A. Moskalik, P. Carson, *et al.*, "3D registration of ultrasound with histology in the prostate," in *Ultrasonics Symposium 2001*, pp. 1397–1400, IEEE, 2001.
7. B. Porter *et al.*, "Histology and ultrasound fusion of excised prostate tissue using surface registration," in *Ultrasonics Symposium 2001*, pp. 1473–1476, IEEE, 2001.
8. Z. Lee, D. B. Sodee, M. Resnick, and G. T. Maclennan, "Multimodal and three-dimensional imaging of prostate cancer.," *Comput Med Imaging Graph* **29**, pp. 477–486, Sep 2005.

9. D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images.," *IEEE Trans Med Imaging* **18**, pp. 712–721, Aug 1999.
10. D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellyn, and W. Eubank, "Nonrigid multimodality image registration," in *SPIE*, M. Sonka and K. M. Hanson, eds., *Medical Imaging 2001* **4322**(1), pp. 1609–1620, SPIE, 2001.
11. M. Ferrant, A. Nabavi, *et al.*, "Registration of 3-d intraoperative mr images of the brain using a finite-element biomechanical model.," *IEEE Trans Med Imag* **20**, pp. 1384–1397, Dec 2001.
12. J. Chappelow, A. Madabhushi, *et al.*, "A combined feature ensemble based mutual information scheme for robust inter-modal, inter-protocol image registration," in *International Symposium on Biomedical Imaging*, IEEE, 2007.
13. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey.," *IEEE Trans Med Imaging* **22**, pp. 986–1004, Aug 2003.
14. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information.," *IEEE Trans Med Imaging* **16**, pp. 187–198, Apr 1997.
15. C. Studholme, D. L. Hill, and D. J. Hawkes, "Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures.," *Med Phys* **24**, pp. 25–35, Jan 1997.
16. H. Matsuda, "Physical nature of higher-order mutual information: Intrinsic correlations and frustration," *Phys. Rev. E* **62**, pp. 3096–3102, Sep 2000.
17. F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(6), pp. 567–585, 1989.
18. S. y. Lu, J. Hernandez, *et al.*, "Texture segmentation by clustering of gabor feature vectors," in *International Joint Conference on Neural Networks, 1991*, pp. 683–688, 1991.
19. J. F. Cardoso, "High-order contrasts for independent component analysis.," *Neural Comp* **11**, pp. 157–192, Jan 1999.
20. H. Stögbauer, A. Kraskov, S. A. Astakhov, and P. Grassberger, "Least-dependent-component analysis based on mutual information.," *Phys Rev E Stat Nonlin Soft Matter Phys* **70**, p. 066123, Dec 2004.