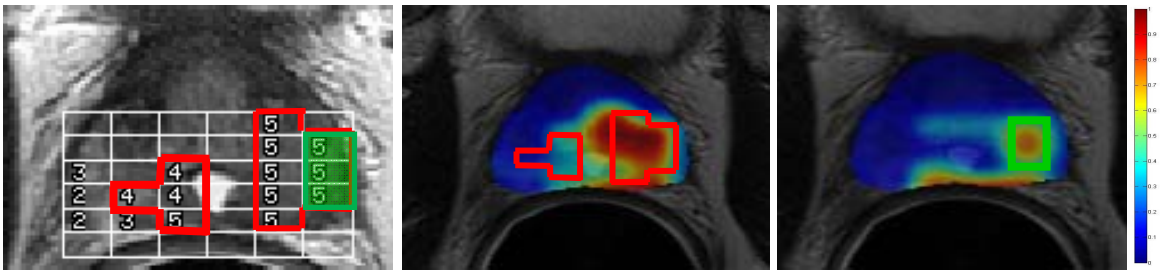
Computerized quantitative data integration of multi-protocol MRI for identification of high grade prostate cancer in vivo.

P. Tiwari[1], J. Kurhanewicz[2], and A. Madabhushi[1]
[1]Biomedical Engineering, Rutgers University, Piscataway, NJ, United States, [2]Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, United States

Background: In this work we present a novel multi-protocol MRI classifier, semi-supervised multi-kernel (SeSMiK), for quantitatively combining features from T2-w magnetic resonance (MR) imaging (T2-w MRI) and MR spectroscopy (MRS) data to identify high grade prostate cancer (CaP) *in vivo*. High grade CaP is known to be correlated with more biologically aggressive prostate cancer; low grade CaP is typically associated with indolent disease. Recently, Shukla et al. [1,2] have demonstrated, qualitatively, that improved differentiability between high and low grade CaP can be obtained using combined T2-w MRI, MRS compared to using the data individually. Choyke et al. [3] have suggested that there is need for quantitative computerized data fusion tools for integrating multi-parametric MRI in order to build classifiers to distinguish between high and low grade prostate cancer *in vivo*. A computerized decision support (CDS) tool that can distinguish high grade from low grade prostate cancer *in vivo* could help identify those patients who might benefit from a "wait and watch policy" as opposed to those who might be better suited to application of more aggressive treatment strategies. The objective of this work is to build a CDS classifier that can distinguish high from low grade prostate cancer *in vivo*, on a per voxel basis, using quantitative integration of T2-w MRI and MRS.

Methods: A total of 19 1.5 Tesla (T) T2-w MRI, MRS studies were obtained prior to radical prostatectomy. All of these studies were biopsy proven prostate cancer patient studies that were clinically referred for a prostate cancer MR staging exam for improved therapeutic selection. T2-weighted MR images were obtained by using the following parameters: 6000/96 (effective), an echo train length of 16, a 3-mm section thickness, no intersection gap, a 14-cm field of view, a 256 × 192 matrix. MRS was acquired as 16 × 8 × 8 phase-encoded spectral arrays (1024 voxels) by using a nominal spectral resolution of 0.24–0.34 cm^3, 1000/130, and a 17-minute acquisition time. An expert spectroscopist manually annotated individual MRS voxels across all studies as firstly (a) CaP/benign, and secondly, (b) as low (≤ 3+4 Gleason score)/high (≥4+3) grade CaP, and these per-voxel MRS annotations were used as a surrogate for ground truth in this study. The 19 studies comprised a total of (a) 573 CaP and 696 benign voxels, and (b) 175 low and 96 high grade CaP voxels. The SeSMiK strategy leverages multi-kernel learning (MKL) [4] and dimensionality reduction (DR) [5] to provide a unified framework for quantitative integration of T2-w MRI and MRS. Kernels capture the pairwise similarities between data points by transforming the input data into a dot product space, and DR homogeneously transforms the high-dimensional data into a low dimensional space by preserving local proximities between the data points. Our SeSMiK strategy comprises of following steps: (1) T2-w MRI texture features based on calculating responses to various gradient filter and gray level co-occurrence operators [6] are obtained along with MRS metabolic peak area features (choline, creatine, citrate, choline+creatine/citrate, choline/creatine), (2) Features extracted in Step (1) from T2-w MRI and MRS are transformed in a common kernel framework K_{MRI} and K_{MRS} respectively, to capture the pairwise similarities between data points across each modality, (3) a weighted combination of pairwise similarities from individual kernels is then obtained as $\hat{K} = \beta K_{MRI} + (1-\beta)K_{MRS}$, where β is the weight assigned to T2-w MRI modality and, (4) the combined kernel $\hat{K}$ obtained in Step (3) is high dimensional and thus cannot directly be used for classification. Hence a low dimensional representation of $\hat{K}$ is obtained using a semi-supervised modification of a popular non-linear DR scheme called graph embedding (SSAGE) [5]. SSAGE incorporates partial labels to provide a better low dimensional representation of $\hat{K}$and hence improved separation between low and high grade prostate cancer.

Results: A probabilistic boosting tree (PBT) classifier which individually evaluated (1) T2-w MRI texture features, (2) MRS metabolic features, and (3) combined low dimensional features obtained via SeSMiK ($\hat{K}$), was employed. Mean AUC for predicting high grade CaP over 25 iterations of randomized cross-validation were found to be, (a) 0.79 ± 0.03 (T2-w MRI), 0.83 ± 0.03 (MRS), and (c) 0.87 ± 0.02 (SeSMiK). SeSMiK also demonstrated higher mean accuracy compared to individual T2-w MRI and MRS.



(a) (b) (c)

Figure 1 (a) CaP ground truth defined by red outline while the green box denotes high grade CaP extent; (b)-(c) show probabilistic heat maps obtained for identifying (b) CaP (identified as red) versus benign (blue) and (c) high grade CaP (red) versus low grade (blue) using SeSMiK on a single T2 slice. Note the high accuracy obtained for CaP and high grade CaP (red) using SeSMiK.

Conclusion: A novel CDS scheme employing SeSMiK is presented for distinguishing between low and high grade CaP *in vivo* via quantitative integration of T2-w MRI and MRS. The classifier was found to have a higher accuracy compared to the individual T2-w MRI, MRS modalities. Our results suggest that the SeSMiK classifier might be able to ultimately identify biologically aggressive CaP *in vivo*.

References: [1] A. Shukla-Dave, et al, *Radiology*, vol. 250, pp. 803-812, 2009. [2] D. L. Langer, et al, *Radiology*, vol. 255, pp. 485-494, 2010. [3] L. Choyke et al. *Nature Reviews*, vol. 6, pp. 191-203, 2009. [4] G. Lanckriet, et al. in Proc. *Pacific Symposium on Biocomputing*, vol. 9, pp. 300-311, 2004. [5] H. Zhao, *Neurocomputing*, vol. 69, pp. 2385-2389, 2006. [6] Viswanath, S. et al, in *Proc. SPIE*, 6915, pp. 69150U–12, 2008.