

COMPUTER-AIDED PROGNOSIS: PREDICTING PATIENT AND DISEASE OUTCOME VIA MULTI-MODAL IMAGE ANALYSIS

Anant Madabhushi, Ajay Basavanahally, Scott Doyle, Shannon Agner, George Lee

Rutgers University
Department of Biomedical Engineering
Piscataway, NJ, USA

ABSTRACT

Computer-aided prognosis (CAP) is a new and exciting complement to the field of computer-aided diagnosis (CAD) and involves developing computerized image analysis and multi-modal data fusion algorithms for helping physicians predict disease outcome and patient survival. At the Laboratory for Computational Imaging and Bioinformatics (LCIB)¹ at Rutgers University we have been developing computerized algorithms for high dimensional data and image analysis for predicting disease outcome from multiple modalities including MRI, digital pathology, and protein expression. Additionally, we have been developing novel data fusion algorithms based on non-linear dimensionality reduction methods (such as Graph Embedding) to quantitatively integrate prognostic information from multiple data sources and modalities. In this paper, we briefly describe 5 representative and ongoing CAP projects at LCIB. These projects include (1) an Image-based Risk Score (IbRiS) algorithm for predicting outcome of ER+ breast cancer patients based on quantitative image analysis of digitized breast cancer biopsy specimens alone, (2) segmenting and determining extent of lymphocytic infiltration (identified as a possible prognostic marker for outcome in Her2+ breast cancers) from digitized histopathology, (3) segmenting and diagnosing highly aggressive triple-negative breast cancers on dynamic contrast enhanced (DCE) MRI, (4) distinguishing patients with different Gleason grades of prostate cancer (grade being known to be correlated to outcome) from digitized needle biopsy specimens, and (5) integrating protein expression measurements obtained from mass spectrometry with quantitative image features derived from digitized histopathology for distinguishing between prostate cancer patients at low and high risk of disease recurrence.

Index Terms— computer-aided prognosis (CAP), breast cancer, prostate cancer, personalized medicine, digital pathology, MRI, data fusion, multi-modal

1. INTRODUCTION

It has always been accepted that cancer is a complex disease which we do not yet fully understand [1]. Predictive, preventive, and personalized medicine (PPP) has the potential to transform clinical practice by decreasing morbidity due to diseases such as cancer by integrating multi-scale, multi-modal, heterogeneous data to determine the probability of an individual contracting certain diseases. In

Funding for this work made possible by the Wallace H. Coulter Foundation, New Jersey Commission on Cancer Research, National Cancer Institute (R01CA136535-01, R21CA127186-01, R03CA128081-01), the Cancer Institute of New Jersey, and Bioimagene Inc.

¹<http://lcib.rutgers.edu/>

the clinic, the same treatment applied to two patients with diseases that look very similar have vastly different outcomes under the same treatment. A part of this difference is undoubtedly patient specific, but a part must also be a result of our limited understanding of the relationship between disease progression and clinical presentation.

An understanding of the interplays of different hierarchies of biological information from proteins, tissue, metabolites, and imaging will provide conceptual insights and practical innovations that will profoundly transform people's lives. There is a consensus among clinicians and researchers that a more quantitative approach, using computerized imaging techniques to better understand tumor morphology, combined with the classification of disease into more meaningful molecular subtypes, will lead to better patient care and more effective therapeutics [2]. With the advent of digital pathology, multi-functional imaging, mass spectrometry, immunohistochemical, and FISH techniques, acquisition of multiple, orthogonal sources of genomic, proteomic, multi-functional imaging, and histologic information for disease characterization is becoming routine at several institutions [1]. Computerized image analysis and high dimensional data fusion methods will likely constitute an important piece of the prognostic tool-set to enable physicians to predict as to which patient may be susceptible to a particular disease and also for predicting disease outcome and survival. These tools will also have important implications in theragnostics; the ability to predict how an individual may react to various treatments, thereby providing (1) guidance for developing customized therapeutic drugs and (2) enable development of preventive treatments for individuals based on their potential health problems. A theragnostic profile that is a synthesis of various biomarker and imaging tests from different levels of the biological hierarchy (genomic, proteomic, metabolic) could be used to characterize an individual patient and her/his drug treatment outcome.

At the Laboratory for Computational Imaging and Bioinformatics (LCIB) at Rutgers University, we have been developing an array of computerized image analysis and high dimensional data analysis, fusion tools for quantitatively integrating molecular features of a tumor (as measured by gene expression profiling or mass spectrometry), results from the imaging of the tumor cellular architecture and microenvironment (as captured in histological imaging), the tumor 3-d tissue architecture and vascularization (as measured by dynamic contrast enhanced (DCE) MRI), and its metabolic features (as seen by metabolic or functional imaging modalities such as Magnetic Resonance Spectroscopy (MRS)). In this paper, we briefly describe 5 representative and ongoing CAP projects at LCIB in the context of breast and prostate cancer and involving digitized histopathology, DCE-MRI, and protein expression measurements obtained via mass spectrometry.

2. IMAGE-BASED RISK SCORE FOR ER+ BREAST CANCERS

The current *gold standard* for achieving a quantitative and reproducible prognosis in estrogen receptor-positive breast cancers (ER+ BC) is via the Oncotype DX (Genomic Health, Inc.) molecular assay, which produces a Recurrence Score (RS) between 0-100, where a high RS corresponds to a poor outcome and vice versa. In [3], we presented Image-based Risk Score (IbRiS), a novel CAP scheme that uses only quantitatively-derived information (architectural features derived from spatial arrangement of cancer nuclei) from digitized ER+ BC biopsy specimens (Figure 1(a)) to help clinicians predict which ER+ BC patients have more aggressive disease and consequently need adjuvant chemotherapy over and above standard hormonal therapy. The hypothesis behind IbRiS is that quantitative image features can be used to implicitly model tumor grade which is known to be correlated with outcome in ER+ BCs.

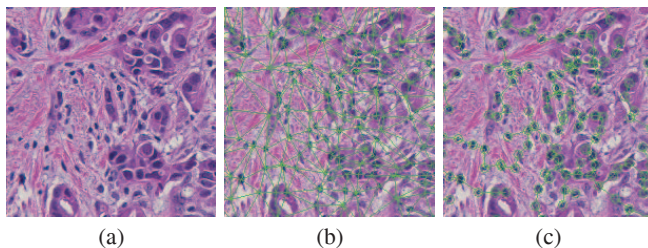


Fig. 1. Nuclear centroids from an (a) ER+ BC histopathology image are used to construct associated (b) Delaunay Triangulation and (c) Minimum Spanning Tree graphs. A total of 12 architectural features are extracted from these graphs and used to quantitatively model phenotypic appearance and hence implicitly the grade of ER+ BC biopsy specimen.

Using the nuclear centroids as vertices, two graphs (Delaunay Triangulation (DT) (Figure 1(b)) and Minimum Spanning Tree (MST) (Figure 1(c))) are constructed and used to extract a total of 12 architectural features for each image. Graph Embedding [4], a non-parametric, non-linear dimensionality reduction technique is employed to project the features onto a reduced 3D space while simultaneously preserving object-class relationships. This allows us to observe the discriminability of the architectural features with respect to low and high RS on a smooth, continuous manifold (Figure 2(a)). The 3D embedding is subsequently unraveled into a normalized 1D IbRiS scale (Figure 2(b)). With a large enough cohort of annotated data, prognostic thresholds θ_1 and θ_2 could be learnt and employed for making prognostic predictions of outcome.

The separation between samples with high and low RS (Figure 2(a)) is reflected quantitatively by the classification accuracy $> 84\%$ [3] of a support vector machine (SVM) classifier. Furthermore, by re-labeling the samples into three classes of low, intermediate, and high RS (Figure 2(b)) we are able to qualitatively confirm that the variations in phenotype described by the architectural features are truly representative of the underlying differences in genotype that affect disease outcome.

3. LYMPHOCYTIC INFILTRATION AND OUTCOME IN HER2+ BREAST CANCERS

The identification of phenotypic changes in BC histopathology with respect to corresponding molecular changes is of significant clin-

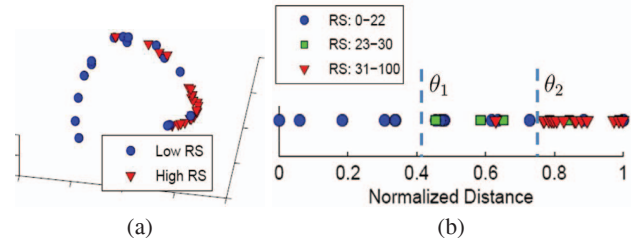


Fig. 2. 37 ER+ histopathology images are plotted in (a) a 3D Graph Embedding space created by reducing the 12 architectural features. The embedding is linearized into (b) the 1D IbRiS scale. Overlaying Recurrence Score labels allows us to identify prognostic thresholds θ_1 and θ_2 for distinguishing poor and good outcome ER+ BC's.

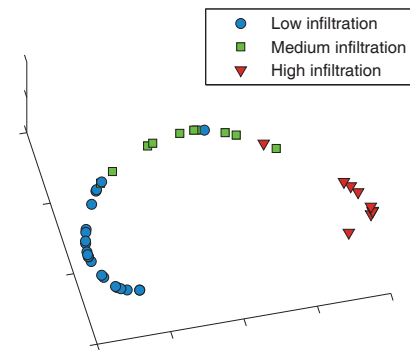


Fig. 3. 41 HER2+ histopathology images plotted in the reduced 3D space obtained by applying Graph Embedding to the original 50 feature (dimensional) space. The manifold of quantitative image features shows a smooth continuum between low, medium, and high levels of LI; ground truth labels were determined by an expert oncologist.

ical importance in predicting BC outcome. One such example is the presence of lymphocytic infiltration (LI) in BC histopathology, which has been correlated with nodal metastasis and distant recurrence in human epidermal growth factor amplified (HER2+) breast cancers. In [5], we introduced a computerized image analysis system for detecting and grading the extent of LI in a digitized HER2+ BC biopsy image. The methodology comprised a region-growing scheme to automatically segment all nuclei (lymphocytic and non-lymphocytic) within the image. The segmentation was then refined via Maximum *a Posteriori* estimation, which utilizes (1) size and intensity information to isolate lymphocytic nuclei and (2) Markov Random Fields to separate clusters of LI from the surrounding baseline level of immune response. The centroids of the resulting lymphocytic nuclei are used to construct graphs (Voronoi Diagram (VD), DT, MST) and a total of 50 architectural features are extracted from each histopathology image. The features are reduced to a 3D embedding space via Graph Embedding. Figure 3 shows this smooth LI manifold, which reveals a continuum between low, intermediate, and high levels of LI. The extent and grade of LI so determined could be converted into a prognostic score (analogous to IbRiS for ER+ BCs and shown in Figure 2(b)) for predicting likelihood of tumor metastasis and distant recurrence.

In conjunction with the architectural features, a SVM classifier was able to successfully distinguish samples with different levels of LI extent at > 90% classification accuracy [5].

4. IDENTIFYING TRIPLE-NEGATIVE BREAST CANCERS ON DCE-MRI

Triple negative (TN) breast cancers lack expression of all 3 major receptors (ER, progesterone (PR), HER2), which are targets of specific therapies. It is important to identify TN cancers quickly since they have an extremely aggressive clinical course and poor prognosis.

The two main lesion descriptors typically used by radiologists when evaluating a breast MRI are the lesion shape (morphology) and the contrast enhancement kinetics. Although breast MRI detects lesions with high sensitivity, the so called BIRADS based descriptors used by radiologists are still somewhat qualitative and thus have low specificity, resulting in high inter-observer variability. Furthermore, since it has been shown that morphology is most likely not a primary descriptor of TN breast cancers, it is extremely important to explore other features such as contrast-enhancement patterns that arise in dynamic contrast enhanced (DCE) MRI images [6]. One

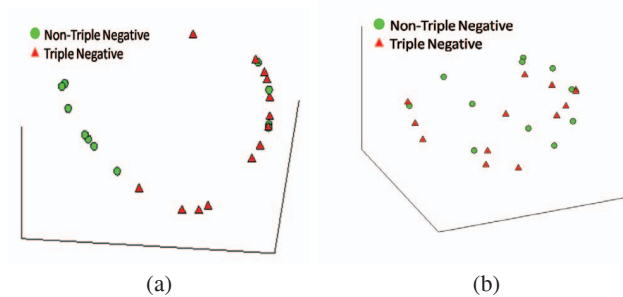


Fig. 4. By reducing image features for (a) the kinetic texture, contrast entropy and (b) signal intensity into a 3D Graph Embedding space, it becomes possible to distinguish between TN and non-TN breast lesions, a task that is difficult for even a trained radiologist.

DCE-MRI slice from each of 24 invasive breast cancers (13 TN, 11 non-TN) were analyzed according to morphology, kinetic signal intensity, static texture, and kinetic texture features. The performance of each feature was analyzed qualitatively using graph embedding and quantitatively using a SVM classifier [6].

The SVM classifier showed that kinetic texture provided the best classification with accuracy of 92%, sensitivity of 92%, and specificity of 91%. Graph embedding on the coefficients fitted to the contrast entropy versus time curves, showed excellent separation between the TN and non-TN lesions (Figure 4(a)). Figure 4(b) shows the corresponding plot for signal intensity versus time, which is not quite as effective in separating the TN and non-TN lesions as kinetic texture (Figure 4(a)).

5. AUTOMATED GLEASON GRADING ON PROSTATE CANCER HISTOPATHOLOGY

Prostate cancer (CaP) is diagnosed in over 200,000 people and causes 27,000 deaths in the US annually. However, the five-year survival rate for patients diagnosed at an early stage of tumor development is very high. If CaP is found on a needle biopsy, the

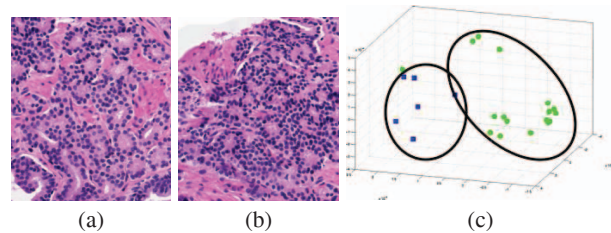


Fig. 5. Gleason grade (a) 3 and (b) 4, prostate cancer biopsy specimens. The results of applying graph embedding to a set of quantitative image descriptors derived from Gleason grade 3 and 4 specimens is shown in (c). Note the excellent separation between grades 3 (green circles) and 4 (blue squares) in (c).

tumor is then assigned a Gleason grade (1-5). Gleason grade 1 tissue is highly differentiated and non-infiltrative while grade 5 is poorly differentiated and highly infiltrative. Gleason grading is predominantly based on tissue architecture (spatial arrangement of nuclei and glands) and tumor morphology (shape and size of glands and nuclei). As tissue regions transform from (a) benign to malignant, and (b) tumor regions transform from lower to higher grades, the architecture and morphology of the images undergo significant changes: nuclear proliferation and infiltration increase, glands in the prostate tissue become smaller, circular, and uniform, and the overall texture of the tissue is altered. Since Gleason grade is known to be strongly correlated to disease outcome, accurately distinguishing between different Gleason grades is critical for making treatment decisions. While pathologists are able to reliably distinguish between low and high Gleason grades (1 and 5), there is a great deal more inter-, and intra-observer variability when it comes to distinguishing intermediate Gleason 3 and 4 patterns (see Figures 5(a), (b)).

At LCIB, we have developed a CAP system that employs morphological, architectural, and textural features derived from prostate needle biopsy specimens [7] to distinguish intermediate Gleason patterns. These features include information traditionally used in the Gleason grading paradigm (morphology and nuclear density) as well as features not considered by pathologists (such as second-order co-adjacency and global texture features). By employing these features in conjunction with a SVM classifier, we were able to distinguish between 40 samples of Gleason grades 3 and 4 with an accuracy of 96.2%. Figure 5 illustrates these results, where each point on the scatter plot represents a CaP biopsy sample (Gleason grade 3 shown with green circles while Gleason grade 4 samples as blue squares). The distance between any two samples is related to their similarity in the original, high dimensional feature space; samples that cluster together have similar feature values and likely to belong to the same Gleason pattern.

6. INTEGRATED PROTEOMIC, HISTOLOGICAL SIGNATURES FOR PREDICTING PROSTATE CANCER RECURRENCE

Following radical prostatectomy (RP), there remains a substantial risk of disease recurrence (estimated at 25-40%). Studies have identified infiltration beyond the surgical margin, and high Gleason score as possible predictors of prostate cancer recurrence. However, owing to inter-observer variability in Gleason grade determination, cancers identified with the same Gleason grade could have significantly different outcomes. Discovery of a predictive biomarker for outcome following RP would allow for therapeutic intervention if the patient

was found to have poor prognosis. Protein expression features of excised prostate tissue may add complementary prognostic information to standard morphologic and architectural features derived from histopathology.

In [8], we attempted to integrate morphological and architectural features quantitatively extracted from digitized excised prostate specimens along with protein expression measurements obtained via electrospray mass spectrometry from the dominant tumor nodule; the idea being to develop an integrated prognostic meta-marker for predicting disease recurrence following RP. To accommodate two widely different modalities (imaging and proteomics), we developed the Generalized Fusion Framework (GFF) for homogeneously representing each of the data types in a normalized, and dimensionality compatible meta-space representation prior to classification in the fused space.

Greater separation between prostate cancer recurrence (red squares) and non-recurrence (green circles) cases was observed in the combined morphologic, architectural, proteomic space (Figure 6(g)) compared to the individual modality spaces (Figures 6(c)-(e)). These results appear to suggest that inclusion of complementary proteomic measurements with traditional Gleason grading based morphologic and architectural measurements may allow for improved prediction of CaP recurrence following RP.

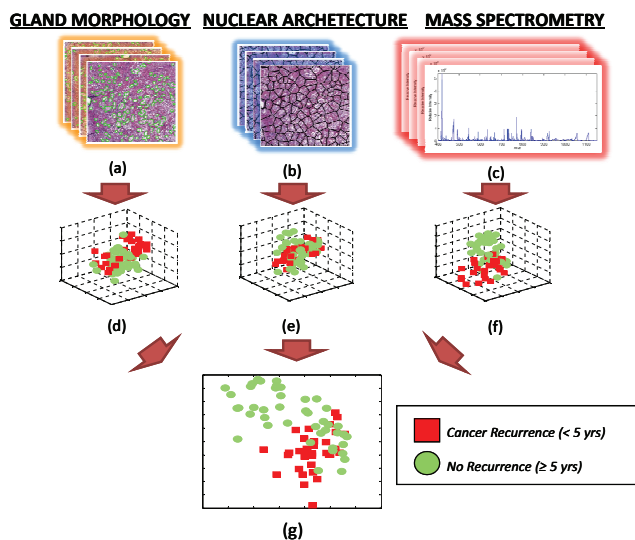


Fig. 6. From excised prostate glands (a) morphological, (b) architectural, and (c) protein expression features were extracted and subsequently reduced to a lower dimensional space via the GFF. The reduced dimensional spaces were then combined. Note the greater separation between CaP recurrence (red squares) and non-recurrence (green circles) cases in the combined morphologic, architectural, proteomic space compared to the individual modality spaces.

7. CONCLUDING REMARKS

In this paper we have highlighted some of our ongoing projects in the newly emerging field of computer-aided prognosis (CAP). While the highlighted projects were focused on prostate and breast cancer outcome prediction, one can easily conceive of similar models being applied to survival prediction for other diseases. Other groups such as Gurcan et al. [9] are applying similar techniques to predict-

ing survival outcome in the context of follicular lymphomas. Further developments in this area will only come about by close and dedicated interactions between computer and imaging scientists, clinicians, oncologists, radiologists, and pathologists.

8. REFERENCES

- [1] A. Madabhushi, "Digital pathology image analysis: opportunities and challenges," in *Imaging in Medicine*, 2009, vol. 1(1), pp. 7–10.
- [2] G. Alexe, J. Monaco, S. Doyle, A. Basavanahally, A. Reddy, M. Seiler, S. Ganesan, G. Bhanot, and A. Madabhushi, "Towards improved cancer diagnosis and prognosis using analysis of gene expression data and computer aided imaging," in *Experimental Biology in Medicine*, 2009, vol. 234(8), pp. 860–79.
- [3] A. Basavanahally, J. Xu, A. Madabhushi, and S. Ganesan, "Computer-aided prognosis of er+ breast cancer histopathology and correlating survival outcome with oncotype dx assay," in *ISBI*, 2009, pp. 851–854.
- [4] A. Madabhushi, J. Shi, M. Rosen, J. Tomaszewski, and M.D. Feldman, "Graph embedding to improve supervised classification and novel class detection: application to prostate cancer," in *MICCAI*, 2005, vol. 8(1), pp. 729–37.
- [5] A. Basavanahally, S. Ganesan, S. Agner, J. Monaco, M. Feldman, J. Tomaszewski, G. Bhanot, and A. Madabhushi, "Computerized image-based detection and grading of lymphocytic infiltration in her2+ breast cancer histopathology," *IEEE Trans. on Biomed. Eng.*, 2009 (accepted).
- [6] S.C. Agner, J. Xu, H. Fatakdawala, S. Ganesan, A. Madabhushi, S. Englander, M. Rosen, K. Thomas, M. Schnall, M. Feldman, and J. Tomaszewski, "Segmentation and classification of triple negative breast cancers using dce-mri," in *ISBI*, 2009, pp. 1227–1230.
- [7] S. Doyle, M. Hwang, K. Shah, A. Madabhushi, M. Feldman, and J. Tomaszewski, "Automated grading of prostate cancer using architectural and textural image features," in *ISBI*, 2007, pp. 1284–1287.
- [8] G. Lee, S. Doyle, J. Monaco, A. Madabhushi, M.D. Feldman, S.R. Master, and J.E. Tomaszewski, "A knowledge representation framework for integration, classification of multi-scale imaging and non-imaging data: Preliminary results in predicting prostate cancer recurrence by fusing mass spectrometry and histology," in *ISBI*, 2009, pp. 77–80.
- [9] L. Cooper, O. Sertel, J. Kong, G. Lozanski, K. Huang, and M. Gurcan, "Feature-based registration of histopathology images with different stains: An application for computerized follicular lymphoma prognosis," in *Comput Methods Programs Biomed*, 2009, vol. 96(3), pp. 182–93.