Integrative Multi-Scale Imaging, Simulation and Precision Cancer Therapy

Joel Saltz MD, PhD

Department of Biomedical Informatics
Stony Brook University

Cancer is a complex multifactorial disease state and the ability to anticipate cancer treatment results and to effectively steer treatment will require synthesis of information across multiple scales from the host to the molecular level. Integration and characterization of molecular and morphological properties of tumors, stroma, metastatic processes, and their interactions with the immune system will also allow understanding of these processes. This has the promise of leading to an in-silico testbed that could be employed to compare treatment options.

While a large variety and number of imaging features are produced and evaluated in imaging studies, at this time there is no integrated framework of methods and tools to enable coordinated curation, management, analysis and assessment of Radiology (Radiomics) and Pathology (Pathomics) imaging feature sets nor to support integrative analysis that combines these feature sets with molecular data to predict outcome and steer treatment.

Development of methods of integrative analysis of multi-scale imaging linked to "omic" characterization is a crucial component of this effort. This effort needs to include integrated analysis of patient data obtained from Radiology, Digital Pathology platforms linked with "omic" information along with analysis of information from novel in vivo and ex vivo imaging platforms. There is a need to bridge the scales, such as to interrelate data obtained through different in vivo and ex-vivo modalities and to interrelate imaging information to spatially mapped "omic" characterizations. Stable and reproducible generation of imaging features such as morphological and texture features of segmented nuclei obtained from Pathology whole slide images along with segmented tumor features obtained from Radiological studies are crucial components of multi-scale analyses.

Integrated multiple scale morphology/molecular tumor/stroma characterization is a challenging exascale application. The exascale nature of this application stems from the confluence of multiple factors. Continuing advances in biomedical imaging and genomic instruments have resulted in an exponential growth of high-resolution imaging and genomic data. Integrated multi-scale analysis of these datasets requires enormous processing and memory capacity. Petabyte-scale raw datasets will not be uncommon in not-too-distant future and will enable such computation intensive work Methods employed in these analyses are pipelines that include iterative image analysis, feature extraction, global optimization and machine learning. While the pipelines are are both data and computationally intensive, the real computing challenges arise from the fact that ensembles of analysis pipelines have

to be systematically executed and their results compared and combined to optimize image feature generation at multiple scales and to associate features with “omics”, outcome and response to treatment. Because there are many ways of characterizing and selecting imaging features and combining this information with information from “omic” platforms, computationally intensive a-posteriori error analysis and stability analysis techniques will be crucial to the success of integrative multi-scale precision medicine methods. These methods and uncertainty quantification workflows need to be executed on tens of trillions of quantitative, spatial, temporal features, generated from Petabyte-scale raw datasets, and on inter-relationships among features within and across scales.

Multi-scale integrative analyses will play crucial roles in drug discovery, the prediction and measurement of response to therapy, the study of the tumor microenvironment, and spatial correlation of imaging phenotypes with genomics. The next step beyond multi-scale integrative characterization is develop principle based prediction methods. We advocate following paths blazed in engineering and physical sciences using extensions of algorithms and High Performance Computing methods developed in those domains. We can now control production in oil fields by combining sensor data (seismic and well pressure data) with simulations, employing porous media models. Our vision is to leverage these same approaches to characterize complex pathophysiology of cancer and to develop methods for adaptive control.