Enabling Scalable Precision Medicine With STAD: A Novel Computational Platform for Early Recurrence-Free Survival Prediction and Biomarker Discovery in Digital Pathology for Gastrointestinal Cancer Patients

Chandrasekhar, Vrishank (School: Lynbrook High School)

Gastrointestinal cancers affect 1.1 million people and are responsible for >875K deaths annually, making them the third leading cause of cancer death worldwide. Late detection of recurrence leads to a 6% survival rate, yet current clinical systems fail to risk-stratify, with a status quo of 53% concordance. This research presents the first low-cost, integrative computational platform to predict the recurrence-free survival rates of gastrointestinal cancer patients two years in advance by leveraging patients' diagnostic tissue slides. First, ~410K histopathology images are preprocessed by training a ResNet-18 model to remove noise, identify cancerous tissue, and generate probabilistic labels. By profiling this information, the model generates datasets of varying tissue types, which are used to train the Cox Proportional Hazards convolutional neural network. The model adaptively evaluates prognostically relevant tissue regions and predicts a risk score for each patient. The identification and prediction models achieved AUROCs of 0.94 and 0.98 across 35 institutions, surpassing all existing clinical techniques. The model's automated analysis of the primary cancer site with limited ground truths offers clinicians the first digital biomarker of recurrence-free survival, sparing patients from overaggressive therapies by providing vital insights into future clinical events. This scalable, affordable, and comprehensive solution can be generalized to other cancer types and diseases and will lead to better clinical outcomes, allow physicians to construct robust, personalized treatment plans earlier, and prevent recurrences from occurring in the first place, saving millions of lives. This work can also be scaled to other medical imaging modalities, further expanding its impact.