A deep learning framework for stratifying cancer patients based on their genomic and interactome profiles

Liang Zhang, Xin Lai

Systems and Network Medicine Lab, Biomedicine Unit, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Human cancer exhibits a diverse landscape of genomic alterations, with certain specific features potentially influencing disease prognosis. The Cancer Genome Atlas (TCGA) provides comprehensive genomic data from cancer patients, facilitating a deeper understanding of the tumor at the molecular level. Leveraging cancer genomics data and gene regulatory networks, we developed a systems medicine approach to characterize the heterogeneity of cancer patients. First, we constructed cancer-specific gene regulatory networks using prior biological knowledge. Second, we identified clusters in the networks and used the information to develop a biologically informed deep learning model. Third, we used the model to learn patterns in the gene expression and DNA methylation data of cancer patients, with the objective of identifying clinically meaningful patient subgroups. We employed this approach to analyze 18 TCGA cancer types, resulting in the identification of patient subgroups exhibiting distinct survival times. Furthermore, we are developing a methodology integrating machine learning techniques with gene regulatory networks to assess the influence of genomic features on the patient stratification outcomes and provide a biological interpretation of the deep learning model. The results show the potential of our approach to enhance the diagnostic utility of pathological examination by integrating tumor morphology with genomics data to predict patients' survival time.