OncoTwin: A Multi-Modal, Treatment-Based Digital Twin for Cancer Recurrence Risk Prediction and Biomarker Discovery

Desai, Dhyana (School: Silver Creek High School)

Breast cancer recurrence rates reach up to 50%, yet existing clinical tools rely on single modality data or repeated medical check-ups. My model, OncoTwin, is a novel recurrence-risk prediction digital twin framework that incorporates four modalities (proteomic, genomic, clinical, and histopathology images) using TCGA-BRCA patients. While previous models utilize predefined gene-signatures, clinically annotated images, and focus on a single treatment, OncoTwin's biomarker discovery and treatment-based architecture reshapes current solutions. In the first molecular module, ~200 RPPA proteins were screened using an Elastic CoxNet Regression model. After applying a LASSO-Cox discordance filter, three strong proteins were identified: PRAS40_pT246, P27|CDKN1B, and DJ1|PARK7. For the second module, 20,155 RNA-sequence genes were filtered using a Log-Rank Pre-filter. A LASSO-Regularized Cox PH model, along with a Random Survival Forest ranking, highlighted 18 strong genes, including C15orf52, IL12B, and PCDHA12. GRAD-CAM features were extracted from histopathology tiles using unsupervised ResNet learning. By combining all of the multi-modal features into a Random Survival Forest, OncoTwin achieves a cross-validated c-index of 0.852. The OncoTwin Interface utilizes PyQt5 to create a patient-specific dashboard to visualize real-time Kaplan-Meier Survival Curves. My model is the first to achieve a high accuracy with multimodal data, unsupervised image analysis without clinical annotation, and treatment-based recurrence risk prediction. OncoTwin's ability to accurately simulate recurrence risk longitudinally surpasses previous approaches' static prediction, setting a high standard for adaptive disease modeling and patient-specific biomarker discovery.

Awards Won:
Third Award of $1,200