

Integrating Multi-Omics, Pathomics, and Radiomics:

A Unified Platform for Breast Cancer Translation

Ori Besen, Tel Aviv University

Breast cancer's biological heterogeneity limits the accuracy of single-modality prognostic models. We present a unified platform integrating seven biological modalities — clinical, mRNA, miRNA, DNA methylation, copy number variation, pathomics, and radiomics — through a 13-microservice Kubernetes architecture for breast cancer survival prediction.

The pathomics pipeline combines TRIDENT tissue segmentation, GAN-based stain normalization, CTransPath feature extraction, and PANTHER prototype encoding into a 16-dimensional slide representation. GAN normalization extended usable WSI coverage by 23% (to 1,052 patients), and the compact 16D representation outperformed 768D CTransPath embeddings (C-index 0.80 vs. 0.74), demonstrating that signal-to-noise optimization can matter more than embedding dimensionality. The radiomics pipeline applies TotalSegmentator across anatomical masks, extracts ~93,000 PyRadiomics features per scan, and distills them through a three-stage funnel to a 1,200-dimensional cross-mask-attention representation.

On TCGA-BRCA (n=1,094), selective fusion of Clinical+DNAm+Pathomics achieved test C-index 0.910 — 17.9% above the clinical baseline — with IBS 0.124 and tertile risk stratification at logrank $p=1.7\times 10^{-6}$. This three-modality combination outperformed full seven-modality fusion, indicating that selective integration beats exhaustive combination. Univariate analysis of PANTHER prototypes identified mucin/myxoid morphology as a top prognostic signal, providing biological validation. A pan-cancer MRI sub-study (n=306, 11 cancer types) achieved C-index 0.836, with prognostic signal emerging from systemic organ features rather than tumor regions — an underexplored direction in radiomics.

The modular architecture enables independent modality development and pan-cancer scaling.