Interpretable Multimodal Machine Learning for Predicting Treatment Response in Breast Cancer

Masoumeh Javanbakhat

1. Abstract

Breast cancer patients with similar clinical and molecular profiles often respond differently to treatment, indicating that current decision-making criteria are insufficient [1]. Histopathology slides contain rich morphological information, yet this signal is rarely integrated with genomic and clinical data in predictive modeling [2]. This project aims to develop interpretable multimodal machine learning models that combine histopathology (H&E), gene expression profiles, and clinical variables to predict treatment response and identify reproducible morphological biomarkers [3]. It aims to create more robust predictive models than those using a single data type and provide insights into the underlying biological mechanisms of cancer progression and treatment resistance. We will (1) learn morphological features predictive of therapeutic outcomes, (2) integrate multimodal data using transformer-based fusion, and (3) validate model-derived biomarkers through collaboration with pathologists and clinicians. The project advances transparent and clinically meaningful AI for precision oncology.

2. Research Vision & Problem Statement

Breast cancer is not a single disease; it is biologically and morphologically heterogeneous [4, 5]. Yet, patients are still frequently treated based on simplified molecular categories, leading to over-treatment for some and ineffective therapy for others. The missing piece lies in the complex morphological patterns visible in histopathology—patterns that reflect tumor organization, immune response, and microenvironmental dynamics—but that are currently interpreted only qualitatively [6].

Recent advances in artificial intelligence and multimodal data integration are redefining this landscape, enabling the combination of histopathological, molecular, and clinical information to better capture the true biological diversity of breast cancer [2, 7]. Such multimodal approaches align with the paradigm of precision oncology, where diagnostic and treatment strategies are tailored to individual patient risk profiles and tumor characteristics, ultimately improving outcome prediction and therapeutic efficacy [4].

My vision is to develop AI systems that augment rather than replace clinical expertise—systems capable of reading morphology at scale, linking it to underlying molecular processes, and producing explanations that clinicians can trust. By designing multimodal, interpretable machine-learning methods that connect tissue morphology to gene expression and treatment outcomes, I aim to contribute to transparent, mechanism-aware precision oncology. The long-term goal is not only improved prediction, but deeper biological understanding to guide therapeutic decision-making.

Research Question: Can interpretable multimodal models linking tissue morphology, molecular activity, and clinical context improve therapy response prediction and reveal reproducible biomarkers in breast cancer?

Hypothesis 1: Predictive Synergy through Multimodality Integrating histopathological morphology with molecular and clinical data will significantly improve the accuracy and robustness of therapy response prediction compared to unimodal models.

Hypothesis 2: Morphology Encodes Reproducible Molecular Signatures Interpretable multimodal models can identify morphological features that are statistically and biologically associated with underlying molecular pathways and can serve as reproducible imaging biomarkers of treatment response.

3. Research Approach and Methodology

3.1 Data and Modalities:

This project will use three primary data modalities to enable multimodal modeling of breast cancer: **Histopathology (H&E) whole-slide images (WSIs)**: Digitized at diagnostic resolution from publicly available breast cancer cohorts (e.g., TCGA-BRCA, CPTAC). WSIs will be tiled into non-overlapping image patches, stain-normalized, and filtered to retain informative tissue regions (epithelium, stroma, immune infiltrates). Morphological features extracted from these tiles will serve as the visual representation of tumor phenotype. **Molecular transcriptomics**: Bulk RNA-sequencing data matched to WSIs will be used to characterize tumor gene expression programs and molecular subtypes. Expression signatures (e.g., immune activity, proliferation, hormone signaling) will provide complementary information about tumor biology and treatment response. **Clinical outcomes and metadata**:

Patient-level variables—including treatment type, receptor status, and survival or therapy response—will serve as supervisory targets and contextual features for multimodal modeling. Integrating these modalities will enable the development of interpretable multimodal models capable of linking morphological and molecular features with clinical outcomes.

3.2 Research Aims and Work Plan

- **Aim 1: Morphological Representation Learning (Years 1–2)** We will train or adapt Vision Transformer-based models (e.g., ViT, CLIP-Histo) combined with attention-based Multiple Instance Learning to derive patient-level morphological embeddings predictive of treatment response. Uncertainty estimation (e.g., Monte Carlo Dropout) will be incorporated to improve robustness.

  Outcome: A quantitative morphological feature space associated with therapeutic outcomes.

- **Aim 2: Multimodal Integration of Histopathology, Genomics, and Clinical Data (Years 2–4)** We will integrate morphology embeddings with gene expression encodings and clinical variables using transformer-based multimodal fusion. Interpretability methods (e.g., attention attribution, concept activation vectors) will be applied to determine how each modality contributes to predictions and what biological structures drive model reasoning.

  Outcome: An interpretable multimodal model outperforming single-modality baselines.

- **Aim 3: Validation of Image-Derived Biomarkers (Years 4–5)** We will identify high-attention tissue regions and extract morphological descriptors (e.g., nuclear shape, stromal density). These features will be correlated with molecular pathways and treatment outcomes in collaboration with pathologists and oncologists. We will quantify discovered biomarkers and assess their reproducibility across independent cohorts to ensure clinical robustness and translational potential.

  Outcome: Reproducible, biologically grounded morphological biomarkers for precision oncology.

4. Collaboration

The Charité, Berlin Institute of Health (BIH), Max Delbrück Center (MDC), and BIFOLD provide a uniquely synergistic research ecosystem that bridges clinical, biological, and computational expertise. Charité and BIH offer extensive clinical and translational resources, including access to well-curated patient cohorts and expert clinical annotation. The MDC contributes leading expertise in mechanistic biology and molecular interpretation, enabling biological validation and pathway analysis. BIFOLD brings cutting-edge capabilities in multimodal and interpretable machine learning, essential for developing and deploying robust AI models. Together, these institutions are connected through shared high-performance computing infrastructure, fostering seamless data integration and collaborative experimentation across domains.

5. Impact and Scalability

This project will generate models that improve treatment response prediction and provide transparent explanations of the underlying decision drivers. The identification of validated morphological biomarkers will support clinical decision-making, individualized therapy selection, and hypothesis generation for mechanism-driven oncology research. By emphasizing transparency and interdisciplinary validation, this work contributes directly to the responsible integration of AI into clinical oncology.

The proposed multimodal architecture is modular and extensible . Each data modality—histopathology, gene expression, and clinical variables—is encoded separately, enabling seamless integration of additional data sources (eg, spatial transcriptomics, proteomics, or MIBI imaging) without redesigning the core model. This structure naturally supports collaboration with clinical and molecular research groups across Charité, BIH, MDC, and BIFOLD during model development and interpretation. The project will adhere to open science and reproducibility standards, with modular, FAIR-compliant pipelines enabling adaptation of the model to new cohorts and cancer subtypes.

6. Risks and Mitigation Strategies

Variability in histopathology slide quality and cohort heterogeneity will be addressed through stain normalization, tile-level filtering, and domain adaptation. Limited sample sizes in treatment subgroups will be mitigated by leveraging foundation-model pretraining, weak supervision, and cross-cohort validation. Integration of heterogeneous modalities will rely on modular encoders and transformer-based fusion architectures that tolerate missing data. To ensure interpretability and clinical trust, pathologists and oncologists will be engaged throughout model development to guide and validate explanations.

References

[1] Emma Nolan, Geoffrey J. Lindeman, and Jane E. Visvader. "Deciphering breast cancer: from biology to the clinic". In: *Cell* 186.8 (2023), pp. 1708–1728. ISSN: 0092-8674.

[2] R.J. Chen et al. "Pan-cancer integrative histology-genomic analysis via multimodal deep learning". In: *Cancer Cell* 40.8 (2022), 865–878.e6.

[3] Pooya Mobadersany et al. "Predicting cancer outcomes from histology and genomics using convolutional networks". In: *Proceedings of the National Academy of Sciences* 115.13 (2018), E2970–E2979.

[4] Xin Xiong et al. "Breast cancer: pathogenesis and treatments". In: *Signal Transduction and Targeted Therapy* 10.1 (2025), p. 49.

[5] Liantao Guo et al. "Breast cancer heterogeneity and its implication in personalized precision therapy". In: *Experimental Hematology Oncology* 12.1 (2023), p. 3.

[6] Rajarsi Gupta et al. "Characterizing Immune Responses in Whole Slide Images of Cancer With Digital Pathology and Pathomics". In: *Current Pathobiology Reports* 8.4 (2020), pp. 133–148. ISSN: 2167-485X.

[7] Óscar Lapuente-Santana, Joan Kant, and Federica Eduati. "Integrating histopathology and transcriptomics for spatial tumor microenvironment profiling in a melanoma case study". In: *npj Precision Oncology* 8.1 (2024), p. 254.