Bulk RNA-seq Guided Multi-modal Detection of Anomalous Regions in Human Cancer via Spatial Transcriptomics

Supplementary Material

6. Details of Baseline Methods

To comprehensively benchmark the performance of our proposed framework, we compare it with a diverse set of state-of-the-art (SOTA) approaches spanning four methodological categories.

Uni-modal histopathology-based methods. This group includes **SimpleNet**[33], which adapts pre-trained feature extractors using a lightweight discriminator against synthetic noise; **THOR**[6], which employs a diffusion model to identify anomalies through deviations from a restored norm; **BIHN**[61], a dual-branch network integrating local and global visual cues; and **AnoDDPM**[30], which leverages denoising diffusion models to localize anomalies via pixel-wise residuals.

Uni-modal ST-based methods. Representative examples include **FICT**[43], a probabilistic generative model jointly capturing gene expression and spatial structure; **Giotto**[12], which constructs spatial networks derived from tissue topology; **SEDR**[50], combining a variational autoencoder with graph convolutional layers to generate unified embeddings; and **SpaGIC**[32], which enhances variational graph autoencoders with adversarial learning for modeling spatial expression patterns.

Multi-modal approaches integrating histology and gene expression. This category comprises **MEATRD**[52], a three-stage fusion architecture leveraging multi-head attention; **STANDS**[51], which integrates spatial GCNs with adversarial domain alignment; **STAIG**[53], based on masked graph reconstruction; **STGIC**[57], enforcing gene embedding consistency through contrastive learning; and **stMMC**[27], which employs parallel graph autoencoders trained using contrastive objectives for multi-modal alignment.

Transfer learning methods. We additionally compare against representative transfer learning frameworks, including **Bulk2Space**[29], a spatial deconvolution model using non-negative matrix factorization; **DARE-GRAM**[37], which aligns domain-specific Gram matrices for regression-based transfer; **IRAD**[54], preserving relational consistency across domains with graph-based adaptation; **LOCIT**[46], performing localized unsupervised source instance selection; and **STEM**[19], which learns a shared encoder to maintain cross-modal consistency between domains.

7. Complexity Analysis

The overall computational complexity of our framework is primarily governed by two core components: the Dynamic Multi-Relational Graph Learning (DMRGL) module and the Optimal Transportation-based Gene Module Alignment (OTGMA) module.

7.1. Complexity of DMRGL Module

The computational cost of the Dynamic Multi-Relational Graph Learning module can be decomposed into two phases:

Dynamic Graph Construction: At the initialization step ($t = 0$), we build the graph using K-Nearest Neighbors (KNN) on spatial coordinates, which has a complexity of $O(N_s \log N_s)$, where N_s denotes the number of spots. For the subsequent refinement steps ($t > 0$), the graph is dynamically updated based on cosine similarity in the gene feature space, requiring $O(N_s^2 D)$ under a naive implementation.

Graph Attention Convolution: In this way, each node computes attention coefficients over its neighbors. Given a graph with average degree k (in our experiments, $k = k_n = k_s$), the complexity per layer is $O(N_s k D^2)$, covering both linear transformations and attention computations across all neighbor pairs. Taking T dynamic refinement steps into account, the overall complexity is dominated by:

$$O\left(T \cdot \left(N_s^2 D + N_s k D^2\right)\right) \tag{14}$$

7.2. Complexity of OTGMA Module

The OTGMA module consists of several computational components:

Gene Similarity Computation: For both the bulk and ST domains, constructing domain-specific gene similarity matrices requires computing $D \times D$ pairwise gene correlations based on domain-specific samples. This results in a total complexity of $O(D^2(N_b + N_s))$, where N_b and N_s denote the numbers of bulk samples and spatial spots, respectively.

Topological Overlap Matrices: Generating topological overlap matrices (TOMs) for K gene modules across the two domains incurs a computational cost of $O(KD^2)$.

Composition Alignment: Computing Wasserstein distances between module Laplacians involves eigenvalue decomposition with worst-case complexity $O(K^2 D^3)$, though low-rank or approximate solvers can reduce this in practice to $O(K^2 D^2)$.

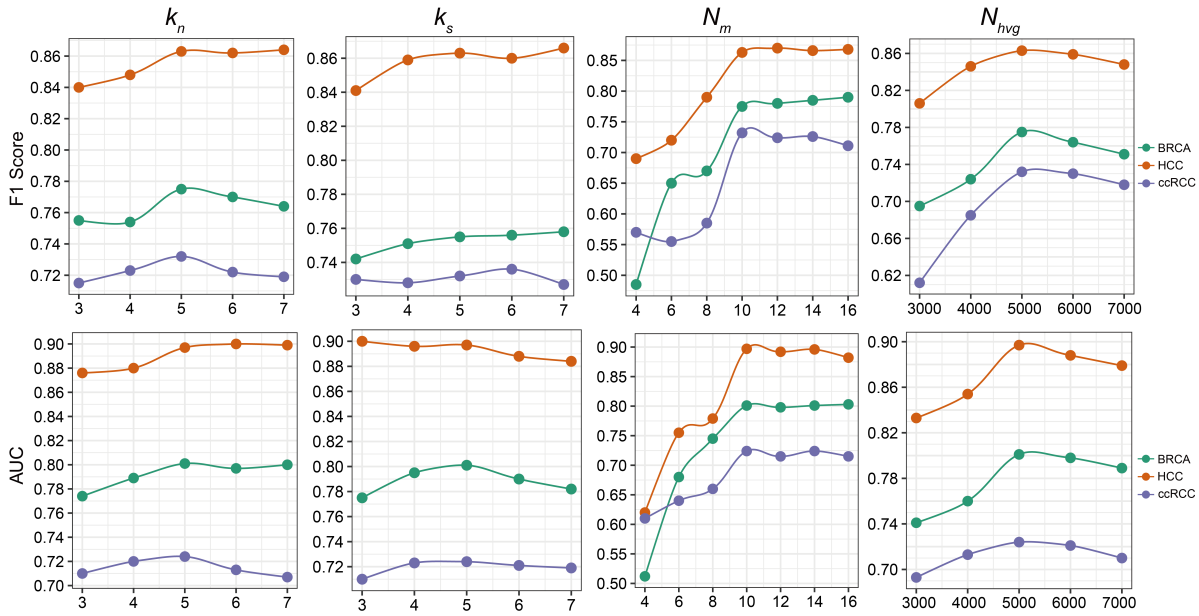

Figure 4. Parameter analysis results. We examine the effects of four parameters: k_n, k_s, N_m, and N_{hvg}.

Functional Alignment: Computing module centroids using the gene embeddings requires $O(KDD_{\text{emb}})$ operations, followed by pairwise module distance computations with complexity $O(K^2 D_{\text{emb}})$.

Optimal Transport: Solving the regularized optimal transport (OT) problem with the Sinkhorn algorithm has complexity $O(K^2 I)$, where I is the number of Sinkhorn iterations.

Combining all components, the total computational complexity of the OTGMA module is:

$$O\Big(D^2(N_b + N_s) + KD^2 + K^2D^3 + KDD_{\text{emb}} + K^2 D_{\text{emb}} + K^2 I\Big) \tag{15}$$

8. Additional Result of Module Alignment

For the HCC cohort, the top bulk-ST module pairs predominantly highlighted pathways involved in metabolic homeostasis, xenobiotic and drug-processing responses, inflammatory activation, and extracellular-matrix remodeling. These functions are consistent with the hallmark metabolic rewiring[15], the impaired detoxification and xenobiotic-metabolism capacity observed in malignant hepatocytes[38], and the inflammation-driven microenvironmental remodeling characteristic of hepatocellular carcinoma[40].

In the case of clear cell renal cell carcinoma (ccRCC), the aligned module pairs revealed strong enrichment in immune regulatory processes, oncogenic signaling programs associated with hypoxia and metabolic reprogramming, and spatially heterogeneous extracellular-matrix remodeling. These pathways correspond closely to the defining biological features of clear cell renal cell carcinoma, including VHL-HIF dysregulation[7] and a highly immune-infiltrated tumor microenvironment[48].

Across all datasets, the aligned module pairs showed consistent enrichment of cancer-relevant pathways. These results indicate that OTGMA achieves biologically meaningful bulk–ST module alignment and effectively reveals tumor-associated molecular processes.

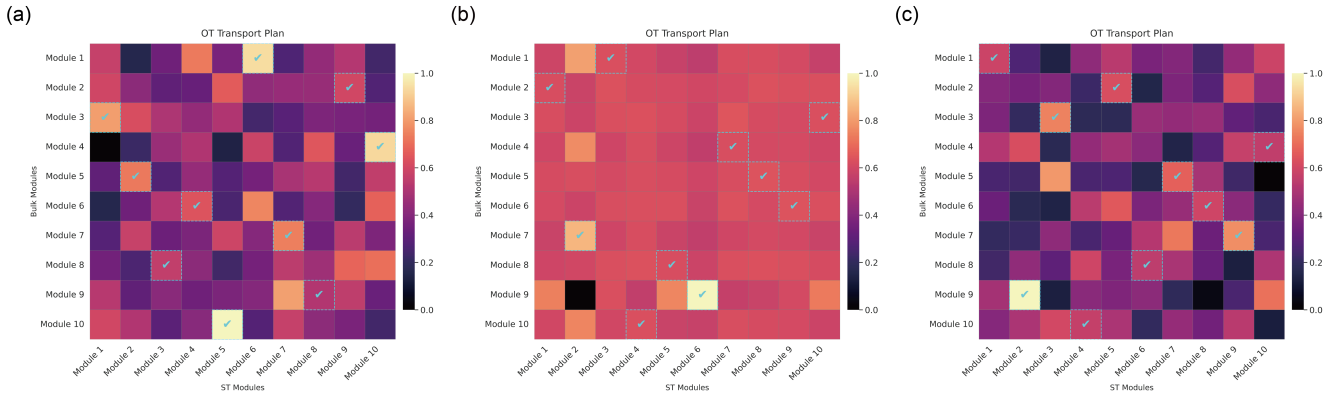

Figure 5. OT transport plan of bulk and ST modules on (a) BRCA, (b) HCC, and (c) ccRCC datasets.

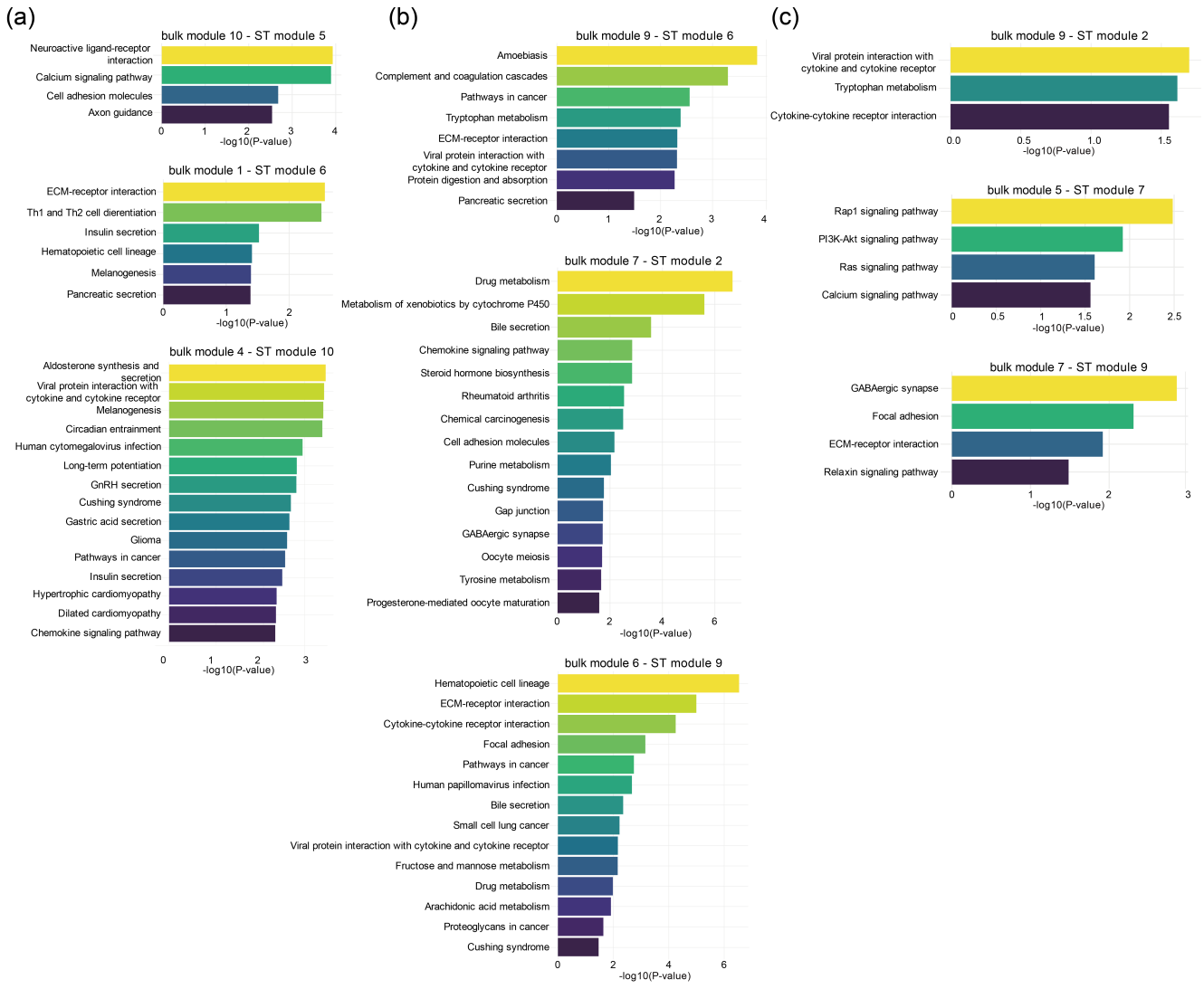

Figure 6. Visualization of KEGG pathways derived from bulk-ST module pairs in the (a) BRCA, (b) HCC, and (c) ccRCC datasets.

Algorithm 1: Cross-domain Gene Module Alignment via Optimal Transport

Input : Domain-specific gene similarity matrices: $\mathbf{A}^m \in \mathbb{R}^{D\times D}$, $m \in \{b, s\}$;
Learnable assignment matrices: $R^m \in \mathbb{R}^{D\times K_m}$.

Output: Transport plan $\mathbf{\Pi} \in \mathbb{R}^{K_b\times K_s}$;
Module-wise TOMs and OT loss.

Step 1: Normalization of Gene Similarity Matrices.

for $m \in \{b, s\}$ **do**
- Symmetrize: $A^m = \frac{1}{2}(A^m + A^{m\top})$;
- Compute degree matrix: $D^m_{ii} = \sum_j A^m_{ij}$;
- Graph normalization: $\hat{A^m} = D^{m-\frac{1}{2}} A^m D^{m-\frac{1}{2}}$.

Step 2: Soft Assignment to Gene Modules.

Initialize R^m with row-stochastic constraints (via softmax).

Step 3: Construction of Topological Overlap Matrices.

for $m \in \{b, s\}$ **do**
- **for** $k = 1, \ldots, K_m$ **do**
  - Compute $\mathrm{TOM}^m_k[i, j] = \hat{A^m}_{ij} \cdot R^m_{ik} \cdot R^m_{jk}$, $1 \le i, j \le D$.

Step 4: Computation of Module Distances.

Compute compositional cost C^{org}_{ij};
Compute functional cost C^{fun}_{ij};
Fuse costs: $C = C^{\mathrm{org}} + C^{\mathrm{fun}}$;
Normalize: $C = C/\max(C)$.

Step 5: Entropic Optimal Transport via Sinkhorn Algorithm.

Define uniform marginals: $a_i = 1/K_b,\ b_j = 1/K_s$;
Gibbs kernel: $\mathbf{K} = \exp(-\mathbf{C}/\varepsilon)$;
Initialize: $u = \mathbf{1},\ v = \mathbf{1}$;
for $t = 1, \ldots, T$ **do**
- $u = a \oslash (\mathbf{K}v)$;
- $v = b \oslash (\mathbf{K}^\top u)$;

Compute transport plan: $\mathbf{\Pi} = \mathrm{diag}(u)\,\mathbf{K}\,\mathrm{diag}(v)$.

Step 6: Optimization and Backpropagation.

Compute OT loss: $\mathcal{L}_{OT} = \langle \mathbf{\Pi}, \mathbf{C} \rangle$;

return $\mathbf{\Pi}, \{\mathrm{TOM}^m_k\}, \mathcal{L}_{OT}$

Table 4. Comparison of model complexity and computational efficiency. We report the number of trainable parameters (MB), inference time (s), and training time per epoch (s) across BRCA datasets.

Method	Trainable Params	Inference Time (s / slide)	Training Time (s / epoch)
SimpleNet	72.1 M	19 ± 0.04	16.65 ± 0.15
THOR	14.2 M	26.2 ± 2.4	12.8 ± 0.9
BIHN	23.5 M	11 ± 0.09	4.1 ± 0.27
AnoDDPM	90 M	40.0 ± 5.0	24.5 ± 2.1
FICT	/	/	/
Giotto	/	/	/
SEDR	10 M	2.0 ± 0.3	8.0 ± 1.5
SpaGIC	8 M	0.8 ± 0.3	2.2 ± 0.7
MEATRD	20 M	2.2 ± 0.6	9.5 ± 1.8
STANDS	38 M	1.3 ± 0.4	8.5 ± 2.0
STAIG	16 M	4.0 ± 0.6	5.2 ± 0.3
STGIC	14 M	0.9 ± 0.3	3.5 ± 0.8
stMMC	9 M	2.0 ± 0.4	5.0 ± 0.6
bulk2space	10 M	1.5 ± 0.5	4.6 ± 1.2
STEM	13 M	0.9 ± 0.1	4.0 ± 1.0
LocIT	11 M	1.1 ± 0.15	2.5 ± 0.5
DARE-GRAM	12 M	1.3 ± 0.05	3.5 ± 0.5
Ours	12.5 M	0.85 ± 0.10	4.0 ± 0.4

Table 5. Summary statistics of spot and gene counts on the (a) BRCA, (b) HCC, and (c) ccRCC datasets.

Dataset	Slice ID	Spot Count	Cancerous Spots	Normal Spots	Gene Count
BRCA	A1	341	303	38	15045
	B1	228	63	165	15109
	C1	167	127	40	15557
	D1	255	139	116	15661
	E1	534	340	194	15701
	F1	659	577	82	14861
	G2	385	160	225	15258
	H1	491	187	304	15029
HCC	HCC-1L	2755	1754	1001	17271
	HCC-2L	4642	2957	1685	16198
	HCC-2P	4661	1358	3303	16198
	HCC-2T	4695	1677	3018	16198
	HCC-3L	4750	3328	1422	17283
	HCC-4L	4092	1980	2112	16020
	HCC-5A	3437	490	2947	16010
	HCC-5B	3942	1916	2026	16010
	HCC-5C	3580	1350	2230	16010
	HCC-5D	4305	1301	3004	16010
ccRCC	GSM5924032	3829	3075	754	17943
	GSM5924033	4975	3130	1845	17943
	GSM5924035	4948	2417	2531	17943
	GSM5924036	4915	3389	1526	17943
	GSM5924037	3585	1574	2011	17943
	GSM5924039	4940	3881	1059	17943
	GSM5924040	4562	3161	1401	17943
	GSM5924041	4359	3186	1173	17943