Supplementary Material

Model Overview

Table 2. Summary of the evaluated representation learning models.

Model	Contrastive	Generative	Shared latent	Private latent
MMVAE [16]	–	✓	✓	–
MMVAE+ [13]	–	✓	✓	✓
MMVAE+sg [11]	–	✓	✓	✓
CLIP [14]	✓	–	✓	–
disSSL [19]	✓	–	✓	✓

Statistical Tables

Table 3. Statistically significant Wilcoxon signed-rank test results after Benjamini–Hochberg FDR correction ($p < 0.05$) across Colon and GBM dataset (HVG, log1p preprocessing). For each slide performance metrics were averaged across repeated runs, and paired slide-level metric values were used as observations in the Wilcoxon signed-rank test. Significance levels refer to the BH-adjusted p-value: $*$ $p < 0.05$; $**$ $p < 0.01$; $***$ $p < 0.001$.

Comparison	Metric	p (BH)	Sig.
Colon			
—	—	—	
GBM			
MMVAE vs MMVAE+sg	PCC (HE→G) 500	0.035	$*$
MMVAE vs MMVAE+	Cosine (G→HE) 200	0.015	$*$
MMVAE vs MMVAE+	Cosine (G→HE) 500	0.019	$*$
CLIP vs disSSL β^{-3}	Cosine (G→HE) 1000	0.019	$*$
CLIP vs disSSL β^{-2}	Cosine (G→HE) 200	0.042	$*$
CLIP vs disSSL β^{-2}	Cosine (G→HE) 1000	0.015	$*$
CLIP vs disSSL β^{-1}	PCC (HE→G) 200	0.028	$*$
CLIP vs disSSL β^{-1}	PCC (HE→G) 500	0.032	$*$
CLIP vs disSSL β^{-1}	Cosine (G→HE) 200	0.015	$*$
CLIP vs disSSL β^{-1}	Cosine (G→HE) 500	0.019	$*$
CLIP vs disSSL β^{-1}	Cosine (G→HE) 1000	0.015	$*$

Training Configurations

Table 4. Training hyperparameters shared across all model configurations.

Hyperparameter	**VAE-based**	**Contrastive**
Optimizer		
Learning rate	1×10^{-4}	1×10^{-4}
Weight decay	1×10^{-5}	—
Optimizer	Adam (AMSGrad)	Adam (AMSGrad)
Training schedule		
Batch size	64	256
Number of epochs	50	50
Architecture		
Latent dimension	32	32
Private latent dim.	16	16
Encoder model layers	[512, 256, 64]	[512, 256, 64]
VAE β-parameter		
β	0.1	—

Sample Overview for HEST Colorectal Cancer Cohort

Table 5. Overview of the retained samples used for model training after filtering of low quality samples and enforcing one sample per patient. HEST ID corresponds to the unique identifier of the sample in the HEST database [7].

Cohort	HEST IDs
Colorectal Cancer	ZEN37, ZEN40, ZEN42, ZEN43, ZEN44, ZEN47, ZEN49

Downstream Task Label Generation Colorectal Cancer Cohort

Pathologist Tissue Annotations. Raw pathologist annotations were obtained from the original publication [18] for each slide. Due to heterogeneity in naming conventions across slides that did not match the categorization criteria, annotations were mapped to a standardized set of tissue region categories. The mapping resolved typographic and formatting variants (normalizing underscores to spaces, converting to lowercase) and merged biologically equivalent labels. Specifically, *connective tissue* classes from slides ZEN37 and ZEN43 were remapped onto corresponding stroma categories based on their immune cell (IC) infiltration grade. *lamina propria*, *squamous epithelium*, and *glandular tissue* were collapsed into *non neoplastic epithelium*. Labels marked as uncertain or appearing in only a single slide were excluded to avoid slide-identity leakage. Each spot was one-hot encoded over the retained categories; unannotated spots were excluded.

Cell Type Proportion Cluster Labels. Cell type count estimations from the original publication were downloaded [18], then normalized to per-spot cell type proportions. To address the compositional nature of these data, proportions were CLR-transformed [6] prior to k-means clustering. Clustering was performed once on the pooled, CLR-transformed proportions from all slides jointly, so that a given cluster label corresponds to the same tissue niche across the dataset without requiring post-hoc label harmonization. The optimal k was selected by silhouette score, yielding four tissue niche clusters in both deconvolution estimations (Belgian & Korean cohort scRNA-seq reference).

Additional Results on Downstream Tasks

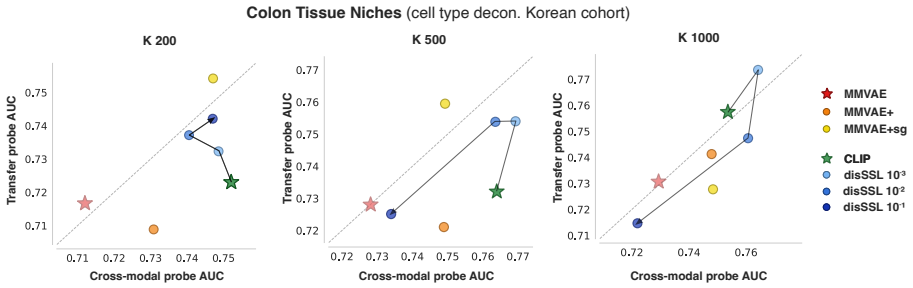

Fig. 3. Downstream task performance for tissue niche classification ($HE \rightarrow G$). AUC prediction performance (mean over five repeated runs with different random seeds) from different gene panel sizes K for tissue niche classification derived from scRNA-seq reference Korean cohort. *Stars* indicate the non-disentangled model and *circles* the disentangled versions. The gray arrow connects contrastive models from CLIP ($\beta = 0$) to increasing disentanglement strength.