Semi-Supervised Contrastive VAE for Disentanglement of Digital Pathology Images — Supplementary Material —

$$\begin{aligned}\mathcal{L}_b(B) = \sum_{j=1}^{m}(& \mathop{\mathbb{E}}_{q_{\phi_z}(z)}[\log p_\theta(B_j|z,s')] - D_{KL}(q_{\phi_z}(z|B_j)||p_b(z))) \\ & - \lambda_1 \cdot L_{MMD}(\hat{q}_{\phi_{s,b}}(s), \delta\{s = s'\})\end{aligned} \tag{3}$$

$$\begin{aligned}\mathcal{L}_c(C_i) = & \mathop{\mathbb{E}}_{q_{\phi_z}(z)q_{\phi_s}(s)}[\log p_\theta(C_i|z,s)] - D_{KL}(q_{\phi_z}(z|C_i)||p_c(z)) \\ & - D_{KL}(q_{\phi_s}(s|C_i)||p_c(s)) - \lambda_2 \mathcal{L}_{MMD}(\hat{q}_{\phi_{z,x}}(z), \hat{q}_{\phi_{z,b}}(z))\end{aligned} \tag{4}$$

$$\mathcal{L}_{M_{recon}} = \mathop{\mathbb{E}}_{q_{\phi_s}(s)}[\log p_{\theta_s}(M|\mu_s)] \tag{5}$$

$$\mathcal{L}_{B_{recon}} = \mathop{\mathbb{E}}_{q_{\phi_z}(z)}[\log p_{\theta_z}(B|\mu_z)] \tag{6}$$

$$\mathcal{L}_{InfoNCE} = -log[\frac{exp(v \cdot v^+/\tau)}{exp(v \cdot v^+/\tau) + \sum_{n=1}^{N} exp(v \cdot v_n^-/\tau)}] \tag{7}$$

$$\mathcal{L}_{KL_s} = D_{KL}(q_{\phi_s}(s|x_i)||p_x(z)) \tag{8}$$

$$\mathcal{L}_{GAN}(D,G) = \mathop{\mathbb{E}}_{x \sim p(x)}[logD(x)] + \mathop{\mathbb{E}}_{s \sim q_{\phi_s}(s), z \sim q_{\phi_z}(z)}[log(1 - D(G(s,z)))] \tag{9}$$

Table 3. Ablation Study of disentanglement models and reconstructors on BRCA dataset. B = Patches without cells, C = Patches with cells, H.T. = High TIL, L.T. = Low TIL, SS = Silhouette Score, S = Salient Latent Space, Z = Background Latent Space

Model	B vs C		H.T. vs L.T.	FID
	$SS_S \uparrow$	$SS_Z \downarrow$	$SS_S \uparrow$	
MM-cVAE (ResNet Architecture)	0.28	0.13	0.11	-
MM-cVAE (ResNet Architecture) + L_κ	0.28	0.13	0.11	-
MM-cVAE (ResNet Architecture) + $L_{M_{recon}}$	0.04	0.12	0.04	-
MM-cVAE (ResNet Architecture) + $L_{B_{recon}}$	0.29	0.10	0.13	-
SS-cVAE (One Step Disentanglement)	0.28	0.08	0.05	-
SS-cVAE with VAE as Reconstructor	-	-	-	251.32
SS-cVAE with IDGAN as Reconstructor	0.32	0.10	0.27	214.59

Table 4. Dataset Statistics. C = patches with cells, B = patches without cells, H.T. = patches with High TIL density, L.T. = patches with low TIL density

Dataset	Tiles (1000 × 1000)			Patches (128 × 128)						Patches (128 × 128)					
	Train	Valid	Test	Train		Valid		Test		Train		Valid		Test	
				B	C	B	C	B	C	L.T	H.T	L.T	H.T	L.T	H.T
BRCA	94	9	10	508	508	14	14	78	78	354	354	30	30	115	115
CoNSeP	27	-	14	479	479	70	70	70	70	398	398	50	50	55	55
Mean and Std of Cell Count Per Tiles and Per Patches for BRCA dataset															
	Cell count in Tiles			Cell Count in C						TIL count in H.T.					
	Train	Valid	Test	Train		Valid		Test		Train		Valid		Test	
mean	281	271	355	6		7		7		9		8		11	
std	134	54	173	3		3		4		3		2		3	
Mean and Std of Cell Count Per Tiles and Per Patches for CoNSeP dataset															
	Cell count in Tiles			Cell Count in C						TIL count in H.T.					
	Train	Valid	Test	Train		Valid		Test		Train		Valid		Test	
mean	576	-	627	10		9		9		9		10		9	
std	477	-	325	6		4		6		7		7		6	
Patches for Downstream Task															
Dataset	Train			Valid				Test							
	Tissue	Other	TIL	Tissue		Other		TIL		Tissue		Other		TIL	
CoNSeP	300	949	349	40		40		40		70		491		268	

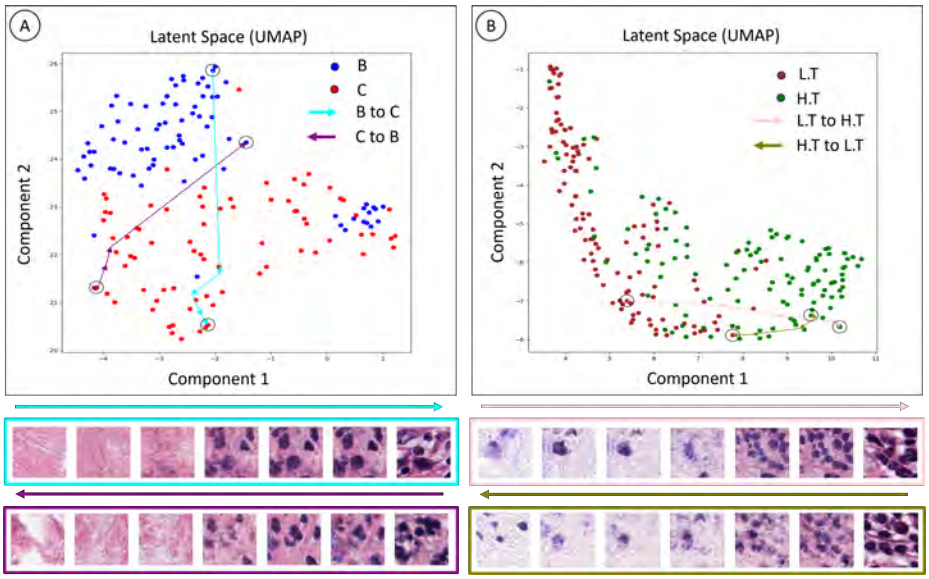

Fig. 6. Latent Space Interpolation. A) Salient latent interpolation from one of the samples from B to C and vice versa visualized with UMAP. B) Salient latent interpolation from one of the samples from $L.T$ to $H.T$ and vice versa. The bottom of both (A) and (B) represents reconstructed images from the interpolated latent. Leftmost and rightmost are original samples.