TopoSlide: Topologically-Informed Histopathology Whole Slide Image Representation Learning – Supplemental Material –

Shahira Abousamra Asmita Sood Sylvia Plevritis
Stanford University, Department of Biomedical Data Science, USA
{sabousamra, asmitas, plevriti}@stanford.edu

A. Implementation Details

TopoSlide is a vision transformer (ViT) [1] configured with 8 layers and 12 multihead attention blocks. It is initialized with weights from timm [2] off-the-shelf pretrained ViT using Momentum Contrast (MoCo) self-supervised learning [3]. The weights assigned to the different components of the loss function are as follows: λ_1 and λ_2 in $\mathcal{L}_{cond-topo}$ (Eq. 1) are set to 0.5, λ_3 and λ_4 in $\mathcal{L}_{local-topo}$ (Eq. 2) are set to 1 and 0.5, respectively, and λ_5 assigned to $\mathcal{L}_{contrast}$ is set to 0.25. AdamW [4] optimizer is used with cosine annealing learning rate scheduler. A warm-up learning rate of 0.01 is used for 5 epochs, then the learning rate is set to 0.001. During training, data augmentation is performed by random sampling of varying-size regions-of-interest (ROIs) from the tissue area with a maximum size of 100×100 patches, where the patch size is 512×512 at 20x magnification. For the loss $\mathcal{L}_{cond-topo}$, 4 patches are randomly sampled from each patch cluster that spans greater than 10% of the slide, in addition to one other patch cluster that has a lower proportion to act as negative sampling. The patch embeddings are used as conditional input to predict the topology and proportion of their respective patch clusters. The persistence diagram for each topological dimension is vectorized into a histogram of persistence to make it more interpretable by the machine learning model. The histogram is binned and normalized using statistics from the training dataset: the normalization is done by dividing each bin by the maximum number of components falling in that bin from within a WSI in the training set.

B. Ablation Study on the Components of the Loss Function

First, we evaluate the contribution of the topological loss. We train the model with only the cluster proportions and the contrastive loss and leave out any topology-based loss terms. The results in Tab. B.1 on the task of the predominant tumor histology retrieval show a big drop in performance when exluding the topology-based loss terms.

Second, we evaluate the contribution of the local topology loss and the contrastive loss when added to the main conditional topological loss. We train TopoSlide on the TCGA-LUAD dataset with different combinations of the loss components and evaluate on the predominant histology retrieval task using the DHMC-LUAD dataset [5]. As shown in Tab. B.2, the performance varies across metrics depending on the combination of the loss components; however, the best overall performance is achieved when all three loss components are combined.

Table B.1. Ablation study on the contribution of the topological loss (Topo. Loss). TopoSlide trained with and without the topological loss terms is evaluated on WSI retrieval performance for DHMC histologic pattern classification. Reporting majority vote classification performance at K=1 and K=5.

		K=1		K=5	
Model	Training Dataset	Bal. Acc	Macro F1	Bal. Acc	Macro F1
TopoSlide (w/o Topo. Loss)	TCGA-LUAD	34.604	32.752	33.782	32.863
TopoSlide	TCGA-LUAD	**52.489**	**51.388**	**46.220**	**49.617**

C. Additional Experimental Details

Due to space limitation, the results for some of the experiments described in the main paper are presented here, as follows.

BRCA Experiments We apply the same experimental structure used for LUAD and train two versions of the model: TopoSlide(TCGA), trained with 50% of TCGA-BRCA (566 slides), and TopoSlide(CPTAC), trained on CPTAC-BRCA (650 slides) to maintain consistent scale with the TCGA-LUAD setting (536 slides). We evaluate performance on three BRCA tasks: histologic subtype retrieval, biomarker prediction with linear probing, and survival prediction using Cox proportional hazards models.

LUAD Gene Mutation Labels Distribution. We indicate the number of positive and negative cases for each label:

Table B.2. Ablation study of loss function, evaluated on whole slide retrieval for histologic pattern classification (C=5) on DHMC-LUAD using majority-vote (MV) prediction over top-K retrieved neighbours. We report MV balanced accuracy and MV macro F-score for $K \in \{1, 3, 5\}$. Best performance is highlighted in bold, and the second-best is underlined.

			K=1		K=3		K=5	
$\mathcal{L}_{cond-topo}$	$\mathcal{L}_{local-topo}$	$\mathcal{L}_{contrast}$	MV Bal. Acc	MV Macro F1	MV Bal. Acc	MV Macro F1	MV Bal. Acc	MV Macro F1
✓	✗	✗	<u>47.268</u>	<u>48.28</u>	35.155	35.859	35.833	36.61
✓	✓	✗	45.046	45.929	37.652	38.126	40.506	41.463
✓	✗	✓	37.714	37.835	**42.549**	**45.165**	<u>41.033</u>	<u>43.255</u>
✓	✓	✓	**52.489**	**51.388**	<u>40.150</u>	<u>42.574</u>	**46.220**	**49.617**

Table C.1. Slide retrieval results for histologic subtype classification (C=2) on TCGA-BRCA using majority-vote (MV) prediction over top-K retrieved neighbours. We report MV balanced accuracy and MV macro F-score for $K \in \{1, 3, 5\}$. Best performance in each model category is highlighted in bold, and the second best is underlined. ** and * indicate the overall best and second-best, respectively. N/A means no slide encoder trained.

	Slide Encoder Pre-training			K=1		K=3		K=5	
Model	Modalities	Dataset	Data Scope	Bal. Acc	Macro F1	Bal. Acc	Macro F1	Bal. Acc	Macro F1
CONCH	Mean Pool	N/A	N/A	**76.431**	**75.489**	**76.855**	**77.922**	**76.566**	**78.320**
Virchow	Mean Pool	N/A	N/A	61.819	62.058	59.709	61.100	57.175	58.308
CTransPath	Mean Pool	N/A	N/A	64.989	65.698	<u>66.918</u>	<u>69.287</u>	<u>66.136</u>	<u>69.370</u>
GigaPath-Tile	Mean Pool	N/A	N/A	<u>67.911</u>	<u>68.427</u>	66.884	68.560	65.933	68.107
TITAN	VLM	Mass 340-K	20 organs	**86.672****	**86.230****	**89.696****	**88.215****	**88.830****	**87.904****
PRISM	VLM	MSKCC, ...	17 organs	82.047*	82.441*	<u>86.246*</u>	<u>85.240*</u>	<u>87.784*</u>	<u>86.302*</u>
CHIEF	VLM	DFCI, TCGA, ...	19 organs	75.469	75.110	79.109	79.649	83.009	82.676
GigaPath	Vision	Providence	31 organs	65.667	66.640	64.917	67.267	63.753	66.279
TITAN-V	Vision	CPTAC-BRCA	BRCA	<u>76.792</u>	<u>77.456</u>	<u>79.333</u>	<u>80.5</u>	<u>80.988</u>	<u>82.357</u>
GigaPath	Vision	CPTAC-BRCA	BRCA	58.543	59.043	58.989	60.302	57.855	59.245
TopoSlide	Vision	CPTAC-BRCA	BRCA	**79.869**	**79.869**	**82.786**	**81.811**	**84.547**	**83.79**
TITAN-V	Vision	TCGA-BRCA	BRCA	<u>76.682</u>	<u>78.93</u>	<u>79.086</u>	<u>80.689</u>	<u>79.856</u>	<u>81.305</u>
GigaPath	Vision	TCGA-BRCA	BRCA	55.861	56.203	53.38	53.506	53.892	53.974
TopoSlide	Vision	TCGA-BRCA	BRCA	**80.381**	**80.258**	**84.055**	**84.263**	**84.428**	**85.14**

CPTAC-LUAD: TP53 (64 positive, 44 negative), EGFR (39 positive, 69 negative), KRAS (35 positive, 73 negative), KEAP1 (11 positive, 97 negative), and STK11. (18 positive, 90 negative). **TCGA-LUAD**: The distribution of mutation labels across 494 patients is: TP53 (280 positive, 214 negative), EGFR (68 positive, 426 negative), KRAS (160 positive, 334 negative), KEAP1 (102 positive, 392 negative), and STK11 (86 positive, 408 negative).

BRCA Biomarker Labels Distribution. We indicate the number of positive and negative cases for each label: **TCGA-BRCA**: ER (810 positive, 238 negative), PR (701 positive, 344 negative), and HER2 (164 positive, 565 negative). **CPTAC-BRCA**: ER (81 positive, 39 negative), PR (68 positive, 47 negative), and HER2 (13 positive, 82 negative).

BRCA Histologic Subtype Retrieval (IDC vs ILC). Breast tumors are primarily categorized into major subtypes: invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC) [6]. We evaluate slide-level retrieval on TCGA-BRCA (647 IDC samples, 183 ILC samples). For each query slide, cosine similarity is computed against all other slides, and the predicted subtype is assigned using majority vote over the top K retrieved neighbors, where K is in $\{1, 3, 5\}$. The results are shown in Tab. C.1 and discussed in the main paper.

Detailed LUAD Gene Mutation Prediction Results Tab. C.3, Tab. C.4: TopoSlide demonstrates strong and comparatively stable performance across key gene mutation prediction tasks across cohorts. When evaluated on TCGA-LUAD, TopoSlide achieves the best TP53 (AUROC: 0.702) and remains competitive for EGFR and STK11 (AUROC: 0.663 and 0.654, respectively). When evaluated on CPTAC-LUAD, TopoSlide trained on TCGA-LUAD outperforms all foundation models on TP53 and STK11, reaching AUROCs of 0.601 and 0.793, respectively, and also achieves the strongest KRAS and KEAP1 performance among the LUAD-specific slide models (0.612 and 0.563). In contrast, although TITAN is

Table C.2. Survival prediction performance using 5-fold cross-validation on TCGA-BRCA (DSS). We report mean and standard deviation of the C-index across folds. Best performance in each model category is highlighted in bold, and the second best is underlined. ** and * indicate the overall best and second best, respectively. N/A means no slide encoder trained.

	Slide Encoder Pre-training			TCGA-BRCA
Model	Modalities	Dataset	Data Scope	DSS
CONCH	Mean Pool	N/A	N/A	$\underline{0.709 \pm 0.071}*$
Virchow	Mean Pool	N/A	N/A	$\mathbf{0.711 \pm 0.047}$**
CTransPath	Mean Pool	N/A	N/A	0.693 ± 0.073
GigaPath-Tile	Mean Pool	N/A	N/A	0.685 ± 0.089
TITAN	VLM	Mass 340-K	20 organs	$\underline{0.697 \pm 0.094}$
PRISM	VLM	MSKCC, ...	17 organs	0.613 ± 0.076
CHIEF	VLM	DFCI, TCGA, ...	19 organs	0.673 ± 0.094
GigaPath	Vision	Providence	31 organs	$\mathbf{0.704 \pm 0.065}$
TITAN-V	Vision	CPTAC-BRCA	BRCA	$\underline{0.667 \pm 0.073}$
GigaPath	Vision	CPTAC-BRCA	BRCA	0.590 ± 0.074
TopoSlide	Vision	CPTAC-BRCA	BRCA	$\mathbf{0.701 \pm 0.072}$
TITAN-V	Vision	TCGA-BRCA	BRCA	$\underline{0.681 \pm 0.113}$
GigaPath	Vision	TCGA-BRCA	BRCA	0.566 ± 0.104
TopoSlide	Vision	TCGA-BRCA	BRCA	$\mathbf{0.690 \pm 0.071}$

one of the strongest foundation models on TCGA-LUAD, its performance drops on CPTAC-LUAD, and its LUAD-specific vision-only variant, TITAN-V has a near chance performance in several settings (AUROC $\in [0.4, 0.5]$ in CPTAC and AUROC $\in [0.45, 0.56]$ in TCGA), suggesting weaker cross-cohort generalization.

Detailed BRCA Biomarker Prediction Results Tab. C.5, Tab. C.6: TopoSlide outperforms most foundation model baselines across BRCA biomarker prediction tasks. On TCGA-BRCA, it achieves the best ER AUROC (0.735) and remains competitive on PR and HER2 (0.621 and 0.541, respectively), compared to the strongest foundation model results of 0.675 for PR and 0.571 for HER2. On CPTAC-BRCA, TopoSlide remains competitive, with the TCGA-BRCA-trained model achieving the strongest HER2 AUROC among all methods (0.638), while the CPTAC-BRCA-trained model attains 0.599, 0.599, and 0.623 AUROC for ER, PR, and HER2, respectively.

Survival Prediction on TCGA-BRCA. We evaluate disease specific survival on TCGA-BRCA using five fold cross validation with a Cox proportional hazards model applied to slide level embeddings. The results are shown in Tab. C.2. TopoSlide(CPTAC) achieves a concordance index of 0.701 ± 0.072, closely matching the strongest baseline model, GigaPath. It outperforms TITAN, PRISM, and CHIEF. TopoSlide(TCGA) also performs strongly with a concordance index of 0.690 ± 0.071. For the CPTAC cohort, all but two patients were censored, which makes survival modeling unreliable; therefore, we do not report survival prediction results for CPTAC-BRCA.

D. Clustering of the Patch Embeddings

TopoSlide learns from the topology of the cluster maps derived from the patch embedding of the Conch model [7]. To confirm that the clusters are biologically meaningful, we asked a board-certified pathologist to annotate sets of patches sampled from the top clusters in the TCGA-LUAD cohort. The patches were selected using two sample approaches:

1. For each cluster, patches were randomly sampled from different slides across the cohort.
2. Within a distinct WSI, patches were sampled from each of the clusters that occupy the largest proportion of the slide.

For the first sampling approach, i.e. cluster patches sampled from different slides, the pathologist was able to assign a label to each cluster that signifies the predominant histologic pattern. Some outliers were noted, as expected. For the seconding approach, the pathologist was also able to assign labels for the predominant histologic pattern for patches pooled from the same cluster in a WSI. High homogeneity within each cluster was noted, with very few outliers. Sample annotations are in Fig. D.1. The pathologist's annotations and notes confirmed that the clusters represent different biological features that are important for the model to learn.

Table C.3. Linear probing results for mutation prediction evaluated on the **TCGA-LUAD cohort**. Evaluation with AUROC and balanced accuracy. Best performance in each model category is highlighted in bold, and the second best is underlined. ** and * indicate the overall best and second-best, respectively. N/A means no slide encoder trained.

	Slide Encoder Pre-training			TP53		EGFR		KRAS		KEAP1		STK11	
Encoder	Modalities	Slide Encoder	Training Scope	AUROC	BalAcc	AUROC	BalAcc	AUROC	BalAcc	AUROC	BalAcc	AUROC	BalAcc
CONCH	Mean Pool	N/A	N/A	**0.684 ± 0.035**	**0.631 ± 0.034**	**0.695 ± 0.017***	**0.638 ± 0.022***	0.556 ± 0.019	0.544 ± 0.018	0.565 ± 0.057	0.533 ± 0.033	**0.660 ± 0.027***	**0.596 ± 0.019**
Virchow	Mean Pool	N/A	N/A	0.633 ± 0.023	0.571 ± 0.033	0.614 ± 0.009	0.546 ± 0.011	0.571 ± 0.016	0.543 ± 0.012	0.538 ± 0.029	0.514 ± 0.023	0.595 ± 0.027	0.538 ± 0.023
CTransPath	Mean Pool	N/A	N/A	0.651 ± 0.020	0.598 ± 0.016	0.632 ± 0.016	0.596 ± 0.012	**0.578 ± 0.019**	**0.555 ± 0.014**	0.561 ± 0.028	0.544 ± 0.021	0.613 ± 0.014	0.558 ± 0.017
GigaPath-Tile	Mean Pool	N/A	N/A	0.624 ± 0.007	0.580 ± 0.023	0.668 ± 0.016	0.584 ± 0.013	0.577 ± 0.026	0.551 ± 0.026	**0.630 ± 0.025***	**0.584 ± 0.048****	0.630 ± 0.027	0.582 ± 0.017
TITAN	VLM	Mass 340-K	20 organs	**0.700 ± 0.015***	**0.652 ± 0.005****	**0.720 ± 0.018****	**0.646 ± 0.031****	0.581 ± 0.008	0.541 ± 0.010	0.477 ± 0.037	0.494 ± 0.007	**0.668 ± 0.030****	**0.630 ± 0.021****
PRISM	VLM	MSKCC, ...	17 organs	0.683 ± 0.011	0.616 ± 0.009	0.694 ± 0.018	0.631 ± 0.012	0.592 ± 0.006*	0.561 ± 0.014*	0.533 ± 0.011	0.514 ± 0.013	0.638 ± 0.021	0.572 ± 0.034
CHIEF	VLM	DFCI, TCGA, ...	19 organs	0.643 ± 0.013	0.588 ± 0.015	0.645 ± 0.029	0.588 ± 0.017	0.565 ± 0.019	0.544 ± 0.009	0.564 ± 0.013	0.554 ± 0.030	0.635 ± 0.010	0.566 ± 0.010
GigaPath	Vision	Providence	31 organs	0.621 ± 0.014	0.566 ± 0.025	0.646 ± 0.013	0.593 ± 0.005	**0.599 ± 0.026****	**0.572 ± 0.027****	**0.632 ± 0.033****	**0.572 ± 0.047***	0.638 ± 0.021	0.599 ± 0.022
TITAN-V	Vision	CPTAC-LUAD	LUAD	0.395 ± 0.019	0.500 ± 0.000	0.507 ± 0.017	0.500 ± 0.000	0.525 ± 0.007	0.500 ± 0.000	0.478 ± 0.035	0.500 ± 0.000	0.538 ± 0.018	0.500 ± 0.000
GigaPath	Vision	CPTAC-LUAD	LUAD	0.631 ± 0.021	0.590 ± 0.025	0.644 ± 0.014	**0.614 ± 0.021**	**0.571 ± 0.008**	**0.546 ± 0.011**	0.507 ± 0.024	0.501 ± 0.014	0.560 ± 0.023	0.539 ± 0.018
TopoSlide	Vision	CPTAC-LUAD	LUAD	**0.690 ± 0.011**	**0.634 ± 0.023**	**0.690 ± 0.013**	0.612 ± 0.019	0.566 ± 0.014	0.543 ± 0.017	**0.525 ± 0.031**	**0.508 ± 0.014**	**0.632 ± 0.023**	**0.544 ± 0.013**
TITAN-V	Vision	TCGA-LUAD	LUAD	0.334 ± 0.013	0.500 ± 0.000	0.403 ± 0.018	0.500 ± 0.000	0.453 ± 0.007	0.500 ± 0.000	0.452 ± 0.009	0.500 ± 0.000	0.446 ± 0.022	0.500 ± 0.000
GigaPath	Vision	TCGA-LUAD	LUAD	0.621 ± 0.013	0.594 ± 0.012	0.614 ± 0.014	0.578 ± 0.012	**0.568 ± 0.019**	**0.548 ± 0.031**	**0.546 ± 0.048**	**0.535 ± 0.024**	0.559 ± 0.015	0.544 ± 0.024
TopoSlide	Vision	TCGA-LUAD	LUAD	**0.702 ± 0.012****	**0.651 ± 0.009***	**0.663 ± 0.014**	**0.604 ± 0.010**	0.527 ± 0.018	0.520 ± 0.007	0.530 ± 0.040	0.527 ± 0.022	**0.654 ± 0.011**	**0.609 ± 0.009***

Table C.4. Linear probing results for mutation prediction tested on the **CPTAC-LUAD cohort**. Evaluation with AUROC and balanced accuracy. Best performance in each model category is highlighted in bold, and the second best is underlined. ** and * indicate the overall best and second-best, respectively. N/A means no slide encoder trained.

	Slide Encoder Pre-training			TP53		EGFR		KRAS		KEAP1		STK11	
Model	Modalities	Dataset	Data Scope	AUROC	BalAcc	AUROC	BalAcc	AUROC	BalAcc	AUROC	BalAcc	AUROC	BalAcc
CONCH	Mean Pool	N/A	N/A	0.524 ± 0.022	0.505 ± 0.019	0.629 ± 0.024	0.567 ± 0.025	**0.620 ± 0.012****	**0.592 ± 0.006****	0.547 ± 0.024	**0.534 ± 0.022**	**0.678 ± 0.017**	**0.625 ± 0.022**
Virchow	Mean Pool	N/A	N/A	**0.571 ± 0.024**	0.533 ± 0.022	0.594 ± 0.025	0.537 ± 0.028	0.592 ± 0.016	0.514 ± 0.027	**0.580 ± 0.056***	0.515 ± 0.036	0.595 ± 0.069	0.533 ± 0.046
CTransPath	Mean Pool	N/A	N/A	0.551 ± 0.008	**0.537 ± 0.008**	0.644 ± 0.024	0.604 ± 0.019*	0.582 ± 0.025	0.559 ± 0.034	0.537 ± 0.032	0.514 ± 0.016	0.653 ± 0.040	0.564 ± 0.038
GigaPath	Mean Pool	N/A	N/A	0.503 ± 0.016	0.519 ± 0.022	**0.734 ± 0.028****	**0.612 ± 0.034****	0.561 ± 0.034	0.534 ± 0.016	0.479 ± 0.033	0.474 ± 0.029	0.619 ± 0.007	0.531 ± 0.024
TITAN	VLM	Mass 340-K	20 organs	0.485 ± 0.030	0.501 ± 0.016	0.633 ± 0.037	0.554 ± 0.029	0.527 ± 0.031	0.511 ± 0.023	0.555 ± 0.036	0.537 ± 0.046	0.679 ± 0.027	0.652 ± 0.027
PRISM	VLM	MSKCC, ...	17 organs	0.561 ± 0.041	**0.553 ± 0.023**	0.641 ± 0.018	**0.587 ± 0.014**	**0.595 ± 0.015**	**0.561 ± 0.006**	**0.638 ± 0.042****	0.551 ± 0.028*	**0.736 ± 0.026***	**0.669 ± 0.014***
CHIEF	VLM	DFCI, TCGA, ...	19 organs	**0.575 ± 0.016**	0.550 ± 0.016	0.630 ± 0.010	**0.587 ± 0.013**	0.545 ± 0.022	0.533 ± 0.020	0.542 ± 0.017	**0.554 ± 0.024****	0.668 ± 0.030	0.621 ± 0.018
GigaPath	Vision	Providence	31 organs	0.525 ± 0.016	0.531 ± 0.018	**0.702 ± 0.024***	0.579 ± 0.029	0.532 ± 0.040	0.505 ± 0.013	0.448 ± 0.042	0.473 ± 0.029	0.699 ± 0.026	0.595 ± 0.017
TITAN-V	Vision	CPTAC-LUAD	LUAD	0.557 ± 0.009	0.504 ± 0.008	0.563 ± 0.062	0.500 ± 0.000	0.468 ± 0.014	0.494 ± 0.010	0.481 ± 0.025	**0.500 ± 0.000**	0.543 ± 0.066	0.500 ± 0.000
GigaPath	Vision	CPTAC-LUAD	LUAD	0.564 ± 0.025	0.554 ± 0.026	**0.637 ± 0.039**	**0.604 ± 0.024***	0.476 ± 0.015	0.499 ± 0.016	**0.499 ± 0.034**	0.485 ± 0.038	0.500 ± 0.009	0.476 ± 0.014
TopoSlide	Vision	CPTAC-LUAD	LUAD	**0.588 ± 0.018***	**0.570 ± 0.017****	0.581 ± 0.013	0.545 ± 0.016	**0.598 ± 0.024**	**0.561 ± 0.030**	0.478 ± 0.016	0.484 ± 0.033	**0.665 ± 0.022**	**0.623 ± 0.020**
TITAN-V	Vision	TCGA-LUAD	LUAD	0.496 ± 0.015	0.500 ± 0.000	0.457 ± 0.010	0.500 ± 0.000	0.453 ± 0.032	0.500 ± 0.000	0.482 ± 0.016	0.501 ± 0.002	0.517 ± 0.029	0.500 ± 0.000
GigaPath	Vision	TCGA-LUAD	LUAD	0.576 ± 0.029	0.551 ± 0.024	**0.636 ± 0.036**	**0.587 ± 0.017**	0.607 ± 0.031	**0.574 ± 0.027***	0.453 ± 0.049	0.488 ± 0.034	0.497 ± 0.040	0.457 ± 0.028
TopoSlide	Vision	TCGA-LUAD	LUAD	**0.601 ± 0.012****	**0.569 ± 0.015***	0.610 ± 0.010	0.564 ± 0.012	**0.612 ± 0.011***	0.550 ± 0.020	**0.563 ± 0.037**	**0.541 ± 0.039**	**0.793 ± 0.024****	**0.713 ± 0.025****

Table C.5. Linear probing results for biomarker prediction evaluated on the **TCGA-BRCA cohort**. Evaluation with AUROC and balanced accuracy. Best performance in each model category is highlighted in bold, and the second best is underlined. ** and * indicate the overall best and second-best, respectively. N/A means no slide encoder trained.

	Slide Encoder Pre-training			ER		PR		HER2	
Model	Modalities	Dataset	Data Scope	AUROC	BalAcc	AUROC	BalAcc	AUROC	BalAcc
CONCH	Mean Pool	N/A	N/A	0.643 ± 0.056	0.581 ± 0.039	0.564 ± 0.035	0.535 ± 0.024	0.530 ± 0.021	**0.524 ± 0.016**
Virchow	Mean Pool	N/A	N/A	0.581 ± 0.044	0.558 ± 0.033	0.559 ± 0.014	0.542 ± 0.016	0.480 ± 0.021	0.493 ± 0.009
CTransPath	Mean Pool	N/A	N/A	**0.661 ± 0.029**	**0.603 ± 0.033***	**0.636 ± 0.022**	**0.594 ± 0.009***	0.500 ± 0.024	0.491 ± 0.017
GigaPath-Tile	Mean Pool	N/A	N/A	0.625 ± 0.023	0.585 ± 0.011	0.592 ± 0.024	0.538 ± 0.016	**0.547 ± 0.016**	0.509 ± 0.014
TITAN	VLM	Mass 340-K	20 organs	0.618 ± 0.043	0.570 ± 0.020	0.617 ± 0.054	0.578 ± 0.026	**0.571 ± 0.026****	0.532 ± 0.015
PRISM	VLM	MSKCC, ...	17 organs	0.579 ± 0.096	0.553 ± 0.064	0.601 ± 0.070	0.565 ± 0.054	0.552 ± 0.023	0.500 ± 0.005
CHIEF	VLM	DFCI, TCGA, ...	19 organs	**0.732 ± 0.021***	**0.659 ± 0.011****	**0.675 ± 0.031****	**0.625 ± 0.020****	0.553 ± 0.026	**0.541 ± 0.020***
GigaPath	Vision	Providence	31 organs	0.645 ± 0.029	0.581 ± 0.017	0.620 ± 0.011	0.521 ± 0.006	0.534 ± 0.008	0.498 ± 0.009
TITAN-V	Vision	CPTAC-BRCA	BRCA	0.327 ± 0.017	0.500 ± 0.000	0.352 ± 0.003	0.500 ± 0.000	0.446 ± 0.017	0.500 ± 0.000
GigaPath	Vision	CPTAC-BRCA	BRCA	0.539 ± 0.021	0.500 ± 0.006	0.499 ± 0.017	0.507 ± 0.003	0.551 ± 0.027	0.506 ± 0.003
TopoSlide	Vision	CPTAC-BRCA	BRCA	**0.704 ± 0.034**	**0.543 ± 0.017**	**0.651 ± 0.034***	**0.529 ± 0.006**	**0.567 ± 0.025***	**0.548 ± 0.020****
TITAN-V	Vision	TCGA-BRCA	BRCA	0.619 ± 0.045	0.500 ± 0.000	0.603 ± 0.020	0.500 ± 0.000	0.423 ± 0.010	0.500 ± 0.000
GigaPath	Vision	TCGA-BRCA	BRCA	0.675 ± 0.012	**0.587 ± 0.033**	0.595 ± 0.019	**0.559 ± 0.023**	0.506 ± 0.505	0.488 ± 0.025
TopoSlide	Vision	TCGA-BRCA	BRCA	**0.735 ± 0.030****	0.547 ± 0.031	**0.621 ± 0.025**	0.519 ± 0.008	**0.541 ± 0.030**	**0.525 ± 0.018**

Table C.6. Linear probing results for biomarker prediction evaluated on the **CPTAC-BRCA cohort**. Evaluation with AUROC and balanced accuracy.Best performance in each model category is highlighted in bold, and the second best is underlined. ** and * indicate the overall best and second-best, respectively. N/A means no slide encoder trained.

	Slide Encoder Pre-training			**ER**		**PR**		**HER2**	
Model	Modalities	Dataset	Data Scope	AUROC	BalAcc	AUROC	BalAcc	AUROC	BalAcc
CONCH	Mean Pool	N/A	N/A	0.606 ± 0.040	0.565 ± 0.034	0.606 ± 0.021	0.547 ± 0.027	**0.580 ± 0.026**	0.541 ± 0.029
Virchow	Mean Pool	N/A	N/A	0.658 ± 0.023	0.554 ± 0.026	$0.661 \pm 0.047*$	**0.585 ± 0.033**	0.458 ± 0.062	0.474 ± 0.017
CTransPath	Mean Pool	N/A	N/A	0.670 ± 0.018	**0.625 ± 0.014***	**0.672 ± 0.024****	0.580 ± 0.025	0.578 ± 0.028	**0.548 ± 0.038**
GigaPath-Tile	Mean Pool	N/A	N/A	**0.675 ± 0.025**	0.584 ± 0.355	0.614 ± 0.027	0.531 ± 0.011	0.485 ± 0.058	0.503 ± 0.035
TITAN	VLM	Mass 340-K	20 organs	$0.681 \pm 0.029*$	0.606 ± 0.020	**0.655 ± 0.029**	**0.608 ± 0.020****	**0.600 ± 0.053**	**0.568 ± 0.043**
PRISM	VLM	MSKCC, ...	17 organs	**0.706 ± 0.029****	**0.650 ± 0.013****	0.599 ± 0.055	0.559 ± 0.032	0.501 ± 0.041	0.484 ± 0.036
CHIEF	VLM	DFCI, TCGA, ...	19 organs	0.653 ± 0.004	0.610 ± 0.015	0.643 ± 0.022	$0.605 \pm 0.027*$	0.561 ± 0.042	0.535 ± 0.037
GigaPath	Vision	Providence	31 organs	0.585 ± 0.025	0.561 ± 0.006	0.526 ± 0.015	0.519 ± 0.006	0.595 ± 0.034	0.563 ± 0.018
TITAN-V	Vision	CPTAC-BRCA	BRCA	0.311 ± 0.009	0.500 ± 0.000	0.332 ± 0.032	0.500 ± 0.000	0.419 ± 0.043	0.500 ± 0.000
GigaPath	Vision	CPTAC-BRCA	BRCA	0.552 ± 0.080	0.531 ± 0.037	0.450 ± 0.035	0.508 ± 0.025	0.605 ± 0.021	0.549 ± 0.012
TopoSlide	Vision	CPTAC-BRCA	BRCA	**0.599 ± 0.010**	**0.557 ± 0.006**	**0.599 ± 0.009**	**0.530 ± 0.003**	**0.623 ± 0.025***	**0.572 ± 0.022***
TITAN-V	Vision	TCGA-BRCA	BRCA	0.459 ± 0.020	0.500 ± 0.000	0.494 ± 0.039	0.500 ± 0.000	0.540 ± 0.022	0.500 ± 0.000
GigaPath	Vision	TCGA-BRCA	BRCA	**0.614 ± 0.004**	**0.574 ± 0.016**	**0.589 ± 0.013**	**0.566 ± 0.010**	0.486 ± 0.027	0.517 ± 0.030
TopoSlide	Vision	TCGA-BRCA	BRCA	0.579 ± 0.036	0.551 ± 0.035	0.540 ± 0.049	0.526 ± 0.025	**0.638 ± 0.033****	**0.587 ± 0.031****

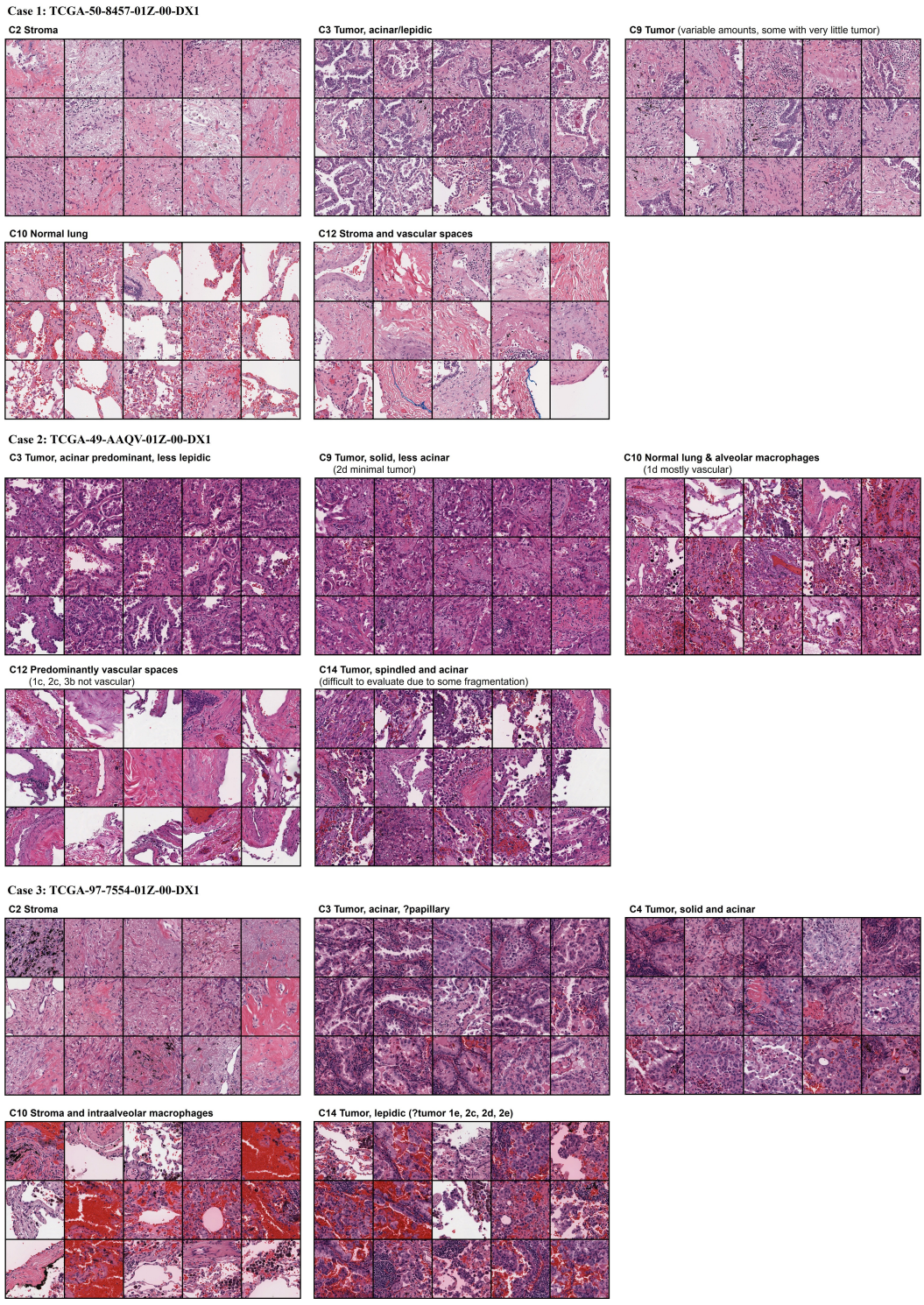

Figure D.1. Case-wise visualization of clusters derived from patch embeddings in the TCGA-LUAD cohort. This figure presents three distinct cases (i.e., Case 1, Case 2, Case 3), each illustrating sample patches from the clusters that occupy the largest proportion of the corresponding slide. Alongside these patches, we include the annotations provided by a board-certified pathologist. Consistent with the previous figure, each WSI and cluster demonstrates a predominant histologic pattern. Notably, the patches in this visualization exhibit greater homogeneity, with reduced noise and fewer outliers compared to earlier analyses [SKP: What is this referring to] . The patches are identified using a grid numbering system, where rows are designated by numeric values and columns by alphabetical letters.

References

[1] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," in 9th International Conference on Learning Representations, ICLR, 2021. 1

[2] R. Wightman, "Pytorch image models (timm)." https://github.com/rwightman/pytorch-image-models, 2019. 1

[3] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9729–9738, 2020. 1

[4] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in 7th International Conference on Learning Representations, ICLR, 2019. 1

[5] J. W. Wei, L. J. Tafe, Y. A. Linnik, L. J. Vaickus, N. Tomita, and S. Hassanpour, "Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks," Scientific reports, vol. 9, no. 1, pp. 1–8, 2019. 1

[6] R. Barroso-Sousa and O. Metzger-Filho, "Differences between invasive lobular and invasive ductal carcinoma of the breast: Results and therapeutic implications," Therapeutic Advances in Medical Oncology, vol. 8, no. 4, p. 261, 2016. 2

[7] M. Y. Lu, B. Chen, D. F. Williamson, R. J. Chen, I. Liang, T. Ding, G. Jaume, I. Odintsov, L. P. Le, G. Gerber, A. V. Parwani, A. Zhang, and F. Mahmood, "A visual-language foundation model for computational pathology," Nature Medicine, vol. 30, no. 3, pp. 863–874, 2024. 3