Semi-supervised Segmentation of Histopathology Images with Noise-Aware Topological Consistency

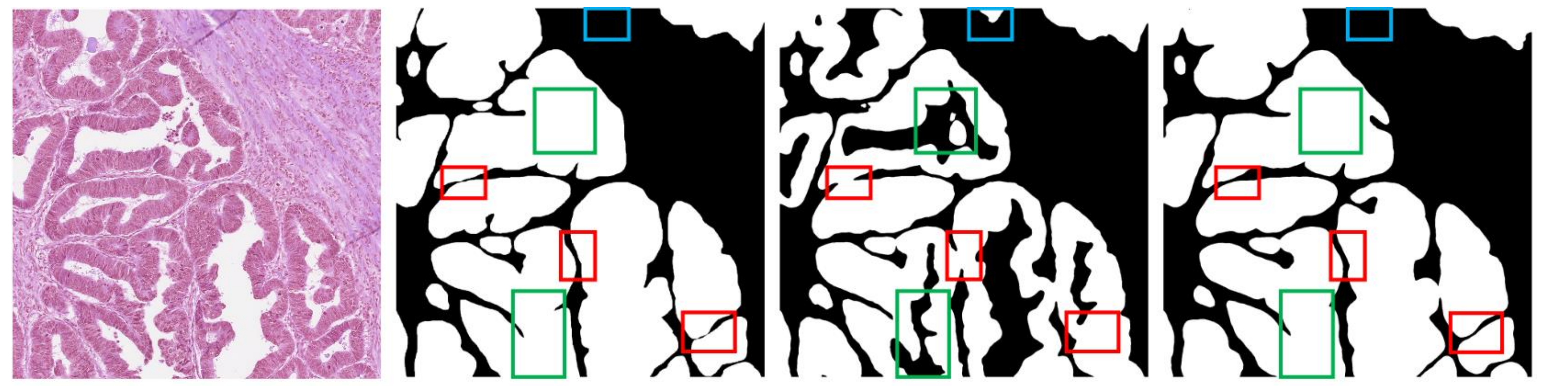
Meilong Xu, Xiaoling Hu, Saumya Gupta, Shahira Abousamra, and Chao Chen

Stony Brook University

Massachusetts General Hospital and Harvard Medical School

Motivations

- In computational pathology, segmenting densely distributed objects like glands and nuclei is crucial for downstream analysis.
- Fully supervised segmentation methods demand a substantial volume of detailed pixel-wise annotations, which is burdensome to obtain.
- Existing semi-supervised learning (SemiSL) methods are often prone to topological errors, such as missing or incorrectly merged/separated glands or nuclei.



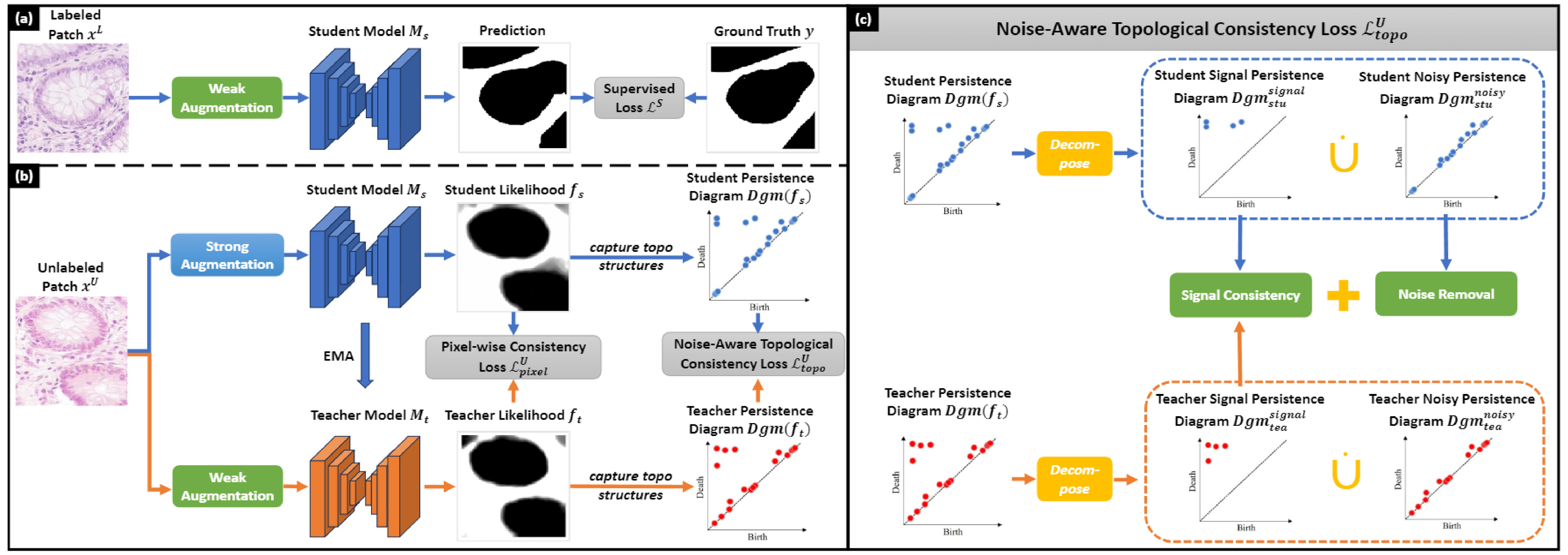
(a) Input (b) GT (c) XNet [62] (d) Ours

Red: prediction errors; Blue: false-positive predictions; Green: false negative holes.

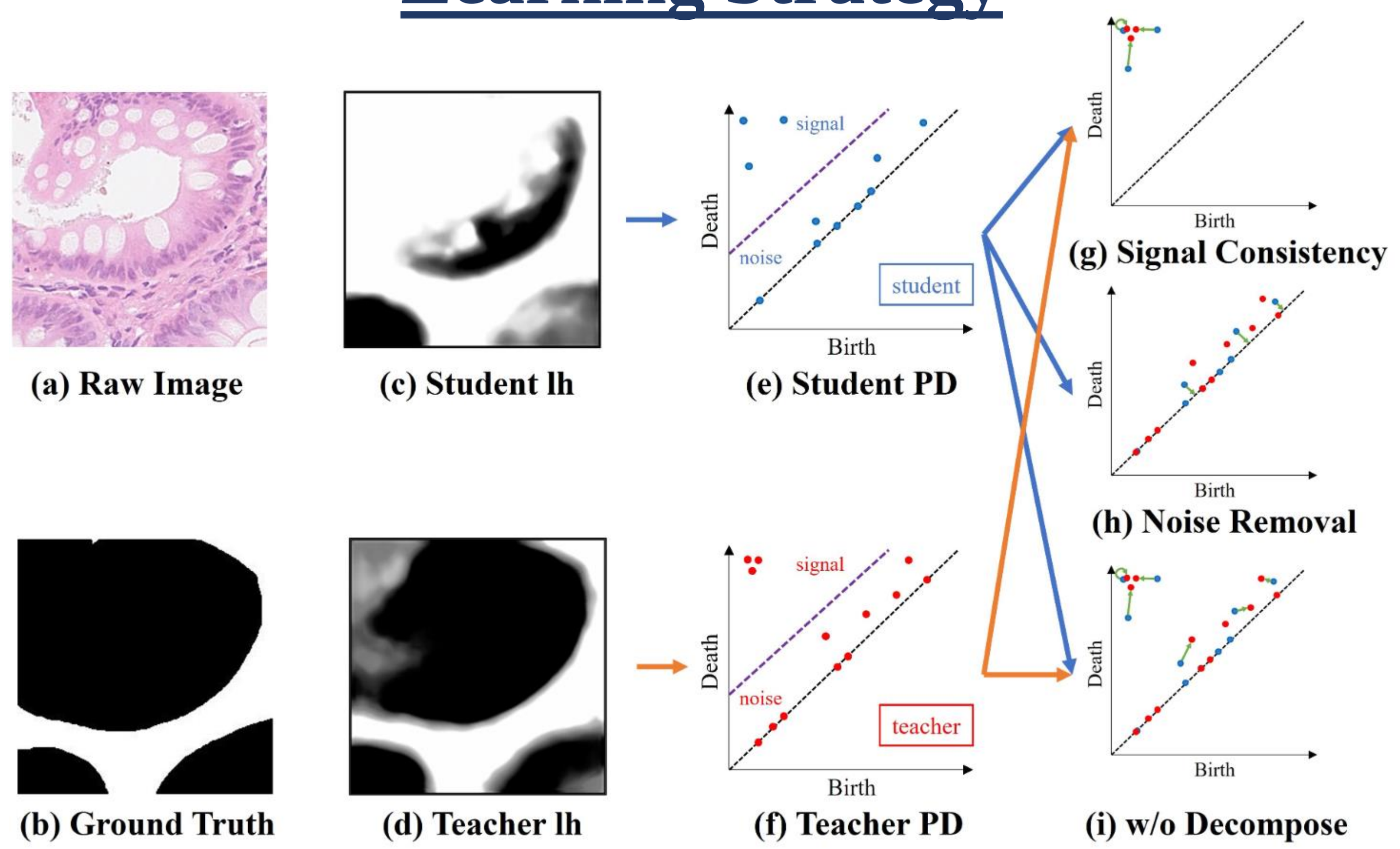
Contributions

- We propose the first topology-aware semi-supervised framework that enforces topological consistency in segmenting densely distributed objects of interest in histopathology images.
- We propose a learning strategy that produces robust topological representations from the noisy topological feature space of the unlabeled images.
- Through extensive experiments on multiple histopathology imaging datasets, we show that our method effectively improves the segmentation quality both pixel- and topology-wise.

Overall Framework



Learning Strategy



(a) Raw Image (c) Student lh (e) Student PD (g) Signal Consistency (h) Noise Removal (b) Ground Truth (d) Teacher lh (f) Teacher PD (i) w/o Decompose

Given a labeled dataset D_L and an unlabeled dataset D_U, we adopt the popular teacher-student framework in SemiSL. This framework contains two networks: student and teacher networks, with identical architecture. **The key intuition is that under different perturbations, the topology of the predictions should be consistent.**

The overall training objective is:

$$\mathcal{L} = \mathcal{L}^S + \mathcal{L}^U$$

To make full use of limited annotations, the supervised loss is formulated as follows:

$$\mathcal{L}^S(D_L, M_s) = \sum_{i=1}^{N_L} [\lambda_1^L \ell_{CE}(M_s(x_i^L), y_i) + \lambda_2^L \ell_{Dice}(M_s(x_i^L), y_i)]$$

We first decompose the persistence diagram into signal and noisy parts based on the persistence:

$$Dgm(f) = Dgm(f)^{signal} \dot{\bigcup} Dgm(f)^{noise}$$
$$Dgm(f)^{signal} = \{p \in Dgm(f) \mid per(p) > \phi\}$$
$$Dgm(f)^{noise} = \{p \in Dgm(f) \mid per(p) \leq \phi\}$$

Then, the designed noise-aware topological consistency loss is as follows:

$$\mathcal{L}^U_{\text{topo}} = \mathcal{L}^U_{\text{topo-cons}} + \mathcal{L}^U_{\text{topo-rem}}$$

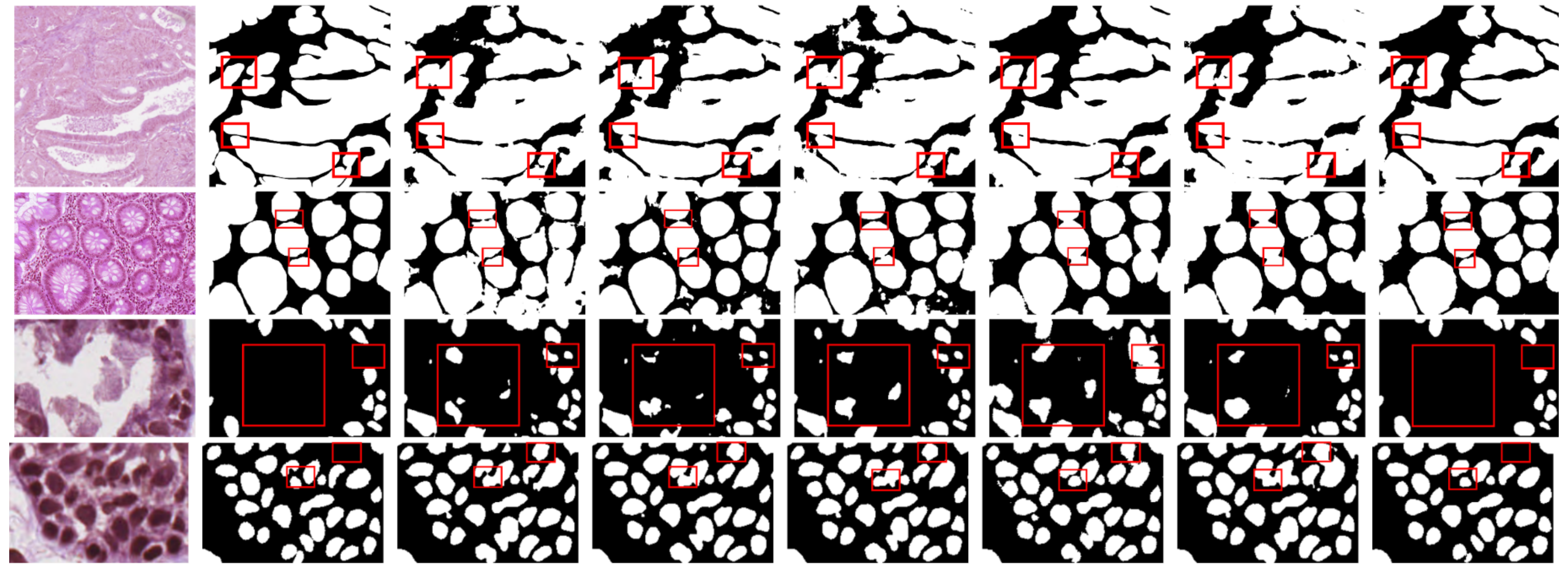
$$\mathcal{L}^U_{\text{topo-cons}} = \sum_{p \in Dgm^{signal}_{stu}} ||p - \gamma^*(p)||^2$$

$$\mathcal{L}^U_{\text{topo-rem}}(f_s) = P_{total}(Dgm^{noise}_{stu}) = \sum_{p \in Dgm^{noise}_{stu}} [f_s(x_p^b) - f_s(x_p^d)]^2$$

Experiments

- Extensive experiments on three datasets are conducted: **Colorectal Adenocarcinoma Gland (CRAG)**, **Gland Segmentation in Colon Histology Images Challenge (GlaS)**, and **Multi-Organ Nuclei Segmentation (MoNuSeg)**.

Qualitative Results



(a) Image (b) GT (c) MT (d) EM (e) UAMT (f) HCE (g) URPC (h) Ours

Quantitative Results

Dataset	Labeled Ratio (%)	Method	Pixel-Wise			Topology-Wise		
			Accuracy ↑	Dice_Obj ↑	IoU ↑	Betti Error ↓	Betti Matching Error ↓	VOI ↓
CRAG	10%	MT [46]	0.862	0.821	0.713	2.238	62.250	0.977
		EM [48]	0.834	0.789	0.688	2.178	80.100	1.027
		UA-MT [58]	0.874	0.837	0.728	1.703	66.450	0.947
		HCE* [25]	0.891	0.862	0.773	1.286	35.530	0.861
		URPC [33]	0.872	0.829	0.728	1.732	74.600	0.883
		XNet [62]	0.895	0.872	0.781	0.578	15.050	0.773
		TopoSemiSeg	**0.905**	**0.884**	**0.798**	**0.227**	**10.475**	**0.758**
	20%	MT [46]	0.887	0.858	0.759	2.603	99.025	0.867
		EM [48]	0.903	0.869	0.776	1.933	75.225	0.798
		UA-MT [58]	0.895	0.859	0.765	1.822	70.850	0.829
		HCE* [25]	0.910	0.881	0.809	0.875	17.400	0.769
		URPC [33]	0.881	0.849	0.744	2.489	99.500	0.912
		XNet [62]	0.907	0.883	0.792	0.422	10.900	0.735
		TopoSemiSeg	**0.912**	**0.898**	**0.820**	**0.226**	**8.575**	**0.709**
	100%	Fully-supervised	0.945	0.928	0.869	0.149	5.650	0.547

Contact

Code: https://github.com/Melon-Xu/TopoSemiSeg

Contact: {meilong.xu, chao.chen.1}@stonybrook.edu