TopoSeg: Topology-Aware Nuclear Instance Segmentation

Hongliang He[1,2*] ✉ Jun Wang[1*] Pengxu Wei[3] Fan Xu[4]
Xiangyang Ji[5] Chang Liu[5] ✉ Jie Chen[1,4] ✉

[1]School of Electronic and Computer Engineering, Peking University, Shenzhen, China
[2]School of Computer Science and Technology, Soochow University, Suzhou, China
[3]Sun Yat-sen University, Guangzhou, China [4]Peng Cheng Laboratory, Shenzhen, China
[5]Department of Automation and BNRist, Tsinghua University, Beijing, China

hehl@pku.edu.cn, junw@stu.pku.edu.cn, weipx3@mail.sysu.edu.cn,
{xuf01,chenj}@pcl.ac.cn, {xyji,liuchang2022}@tsinghua.edu.cn

Abstract

Nuclear instance segmentation has been critical for pathology image analysis in medical science, e.g., cancer diagnosis. Current methods typically adopt pixel-wise optimization for nuclei boundary exploration, where rich structural information could be lost for subsequent quantitative morphology assessment. To address this issue, we develop a topology-aware segmentation approach, termed TopoSeg, which exploits topological structure information to keep the predictions rational, especially in common situations with densely touching and overlapping nucleus instances. Concretely, TopoSeg builds on a topology-aware module (TAM), which encodes dynamic changes of different topology structures within the three-class probability maps (inside, boundary, and background) of the nuclei to persistence barcodes and makes the topology-aware loss function. To efficiently focus on regions with high topological errors, we propose an adaptive topology-aware selection (ATS) strategy to enhance the topology-aware optimization procedure further. Experiments on three nuclear instance segmentation datasets justify the superiority of TopoSeg, which achieves state-of-the-art performance. The code is available at https://github.com/hhlisme/toposeg.

1. Introduction

Pathological image analysis is considered the gold standard for cancer diagnosis, treatment, and prevention, where nuclear instance segmentation serves as the core step. It is because the morphology of nuclei including the shape, appearance, and distribution within pathology images provides interpretable features of diagnostic and prognostic cancer indicators[22].

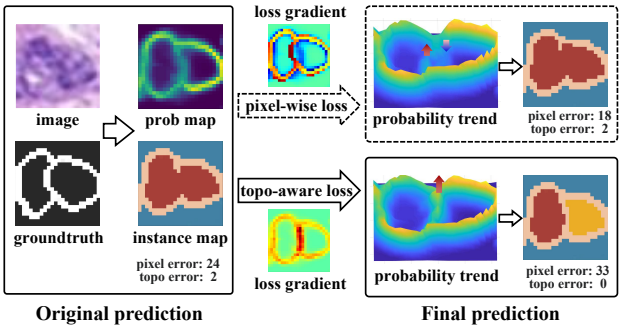

Figure 1. **The schematic illustration of pixel-wise loss and topology-aware loss.** (Left) A patch in the image and its corresponding boundary ground truth, as well as the original boundary segmentation probability map and its corresponding instance prediction. (Top-right) The gradient map was obtained by pixel-wise loss, the 3D map of the boundary probability map and the optimization trend, as well as the final instance prediction. (Bottom-right) The gradient map obtained from topology-aware loss, the 3D image of the boundary probability map and the optimization trend, as well as the final instance prediction.

Compared to general image object segmentation, nuclear instance segmentation faces with some particular challenges: (i) the crowded nuclei distribution leads to clustered overlapping instances. (ii) The blurry boundaries between touching or overlapping nuclei increase the difficulty of distinguishing individual instances. (iii) Large variability in size, shape, texture and intensity of nuclei between different images requires more robust annotation and prediction [18].

To cope with these challenges, several studies have focused on exploring innovative network architectures [12, 23, 25] and incorporating distance maps [3, 9, 19, 21] to enhance segmentation performance. However, existing works typically rely on pixel-wise losses, such as CE/Dice loss,

*Equal contribution ✉ Corresponding Author

which enforce predictions to align with GT pixel-by-pixel yet ignore the crucial structure information of instances. Since treat pixels independently, they lack sensitivity towards high-level structural characteristics, such as the connectivity of boundaries, which are essential for the rational segmentation of individual instances. This inability of pixel-wise loss often leads to unreasonable results such as broken boundaries, bringing about failure to distinguish touching or overlapping instances. This consequently results in undesirable effects on the subsequent quantitative morphology analysis. As depicted in Fig. 1 (top-right), pixel-wise loss focuses on misaligned pixels thus reducing pixel errors. However, due to the lack of perception of structural information, the boundaries between overlapping instances with smaller pixel proportions are not given enough consideration and result in incorrect instance predictions.

To capture structural information of nuclear instances, we propose a novel topology-aware nuclear instance segmentation approach, termed TopoSeg. Specifically, we introduce a topology encoding method to encode topological features of three-class maps which are utilized to separate individual instances. Building upon this, we propose a topology-aware module (TAM), which encodes dynamic changes and persistence of different topology structures within the predicted three-class probability maps to persistence barcodes and develops a topology-aware loss function, integrating topology information into model learning. As shown in Fig. 1 (bottom-right), with topology constraints, the critical boundaries are identified and optimized, resulting in a reasonable segmentation outcome. To further adapt to the task of dense and small instances like nuclei, we introduce an adaptive topology-aware selection module (ATS) to concentrate on images and regions requiring topological optimization. Notably, TopoSeg effectively promotes instance segmentation performance without requiring any changes to the original network structure.

The contributions can be summarized as follows:

- We propose TopoSeg, which provides a new perspective that goes beyond the common pixel-wise optimization paradigm and incorporates meaningful topology constraints for nuclear instance segmentation.
- We present a TAM to encode the persistence of topology structures in probability maps and introduce a topology-aware loss as a further constraint.
- We develop an ATS to identify regions that most require topology optimization, enabling the model to efficiently accommodate segmentation tasks involving small and dense objects.
- We justify the superiority of our TopoSeg on three nuclei datasets with state-of-the-art performance.

2. Related Work

2.1. Nuclei Segmentation

Deep learning has been extensively utilized for nuclear instance segmentation [2, 9, 11, 24]. In 2015, Ronneberger et al. [25] proposed the U-Net model, which has become one of the most fundamental models in medical image segmentation. Raza et al. [24] proposed Micro-Net, which achieves robustness to large internal/external variances in nucleus size by utilizing multi-resolution and weighted loss functions. Qu et al. [23] proposed a full-resolution convolutional neural network (FullNet) with no downsampling operation employed in the network structure to enhance localization accuracy. He et al. [12] proposed a hybrid attention nested U-shaped network (Han-Net) to extract effective feature information from multiple layers.

To take advantage of contour information to differentiate contact/overlapping nuclei, Chen et al. [2] initially proposed incorporating contour information into a multi-level FCN to create a deep contour-aware network for nuclear instance segmentation. Subsequently, Zhou et al. [28] proposed the contour-aware information aggregation network to combine spatial and textural features between nuclei and contours. Additionally, some works [3, 9, 19, 21] introduced distance maps to separate contact/overlapping nuclei. Naylor et al. [21] addressed the issue of segmenting touching nuclei by formulating the segmentation task as a regression task of intra-nuclear distance maps. Graham et al. [9] proposed a network (Hover-Net) for simultaneous nucleus segmentation and classification, which uses the vertical and horizontal distances between a nucleus pixel and its center of mass to separate clusters of nuclei. Moreover, He et al. [11] proposed a centripetal directional network (CDNet) for nuclear instance segmentation, combining directional information into the network.

The above models have achieved some performance improvements in nuclear segmentation, but these models using pixel-wise loss tend to ignore structural information of nuclear boundaries.

2.2. Segmentation Topology

Several topology-based approaches [1, 14, 16, 20, 26] have exploited ways to guarantee the topological accuracy of objects in images. Mosinska et al. [20] considered topological information by constraining the similarity between feature maps and predicted probability maps generated by a pre-trained VGG-19 network. Shit et al. [26] proposed a similarity metric (clDice) to calculate the intersection between segmentation masks and their morphological skeletons, and developed a novel loss function (Soft-clDice) to train the neural network based on the skeleton structure.

In addition, the differentiability of persistent homology [6] allows it to be formulated as a loss function and ap-

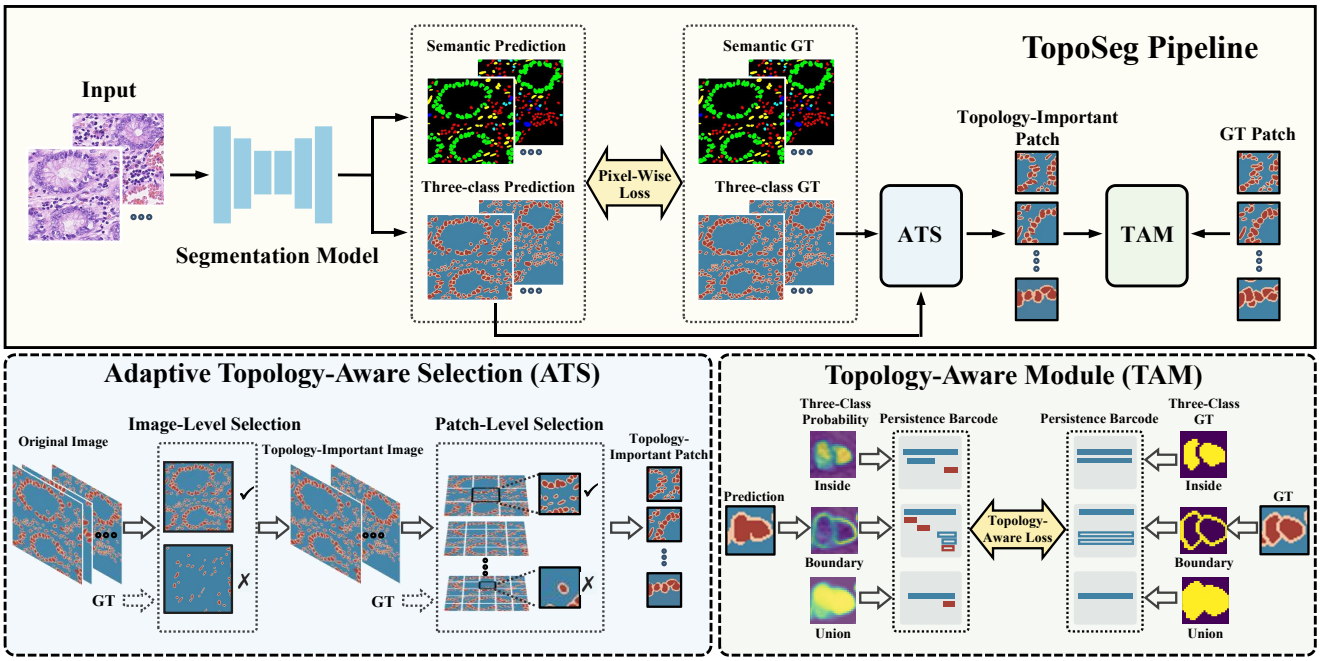

Figure 2. **The overall framework of the proposed TopoSeg.** The network contains two basic prediction branches: the semantic branch and the three-class (inside, boundary, background) branch. In order to be aware of topology, the framework contains two important modules: the adaptive topology-aware selection module and the topology-aware module.

plied to end-to-end deep neural networks. Hu et al. [14] developed a loss function that directly constrains the prediction results and topological GT based on persistent homology.Furthermore, Hu et al. [16] used Discrete Morse Theory [7, 8] to capture the singularities of the gradient vector field of the likelihood function to find all potential skeletons of the object. They also leveraged Discrete Morse Theory and persistent homology to construct a topological representation space and trained a probabilistic model for performing inference tasks in such a topological representation space [15]. In cardiac magnetic resonance (CMR), Clough et al. [1, 5, 4] investigated various methods to couple topological information, where the number of components in the required segmentation is known a priori.

However, the above methods have certain limitations in terms of computational efficiency, topology completeness, and task flexibility. Therefore, we propose a topology-aware model for nuclear instance segmentation, which encodes structural information to construct topological losses, thus improving the instance segmentation performance.

3. Methodology

3.1. Overview

Fig. 2 provides a flowchart of our TopoSeg, which adopts a dual-branch architecture with two types of predictions: semantic prediction of nuclear semantic categories (*e.g.* carcinoma, epithelial, etc.) and three-class prediction of structural categories (inside, boundary, background). The output of three-class branch is sequentially processed through an adaptive topology-aware selection module (ATS) and a topology-aware module (TAM) to efficiently capture the topology information of nuclear instances.

Specifically, to quantify the topological features of an image, we present topology encoding in Sec. 3.2. Topological information contained in three-class images, such as the number of holes and connected components of the boundary, can reflect the number of instances and overlapping characteristics, which is meaningful for reasonable instance segmentation. Therefore, we decompose the three-class image into three channels (inside, boundary, and union) and encode their topological features. This allows the direct representation of the structural correctness, but its inapplicability to predicated probability maps makes it challenging to incorporate into the training process.

To address this issue, we introduce TAM in Sec. 3.3, which extracts topological structures of predicted three-class probability maps into persistence barcode through a persistence encoding method and constrains the topology similarity between GTs and predictions by a topology-aware loss.

To further improve the topology perception of the model and reduce the computational cost, we develop ATS in Sec. 3.4 to adaptively select regions with greater topology importance in both image- and patch-level, which makes the topological constraints more suitable for the task of seg-

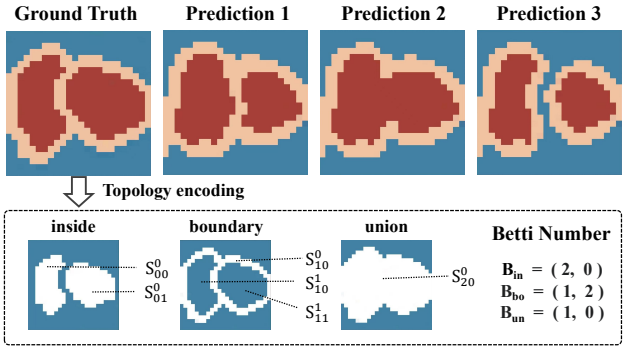

Figure 3. (Top) The ground truth and different prediction results of two overlapping nuclei in the pathological image. (Bottom) An example of topology encoding of ground truth.

menting dense and small instances.

In general, during the training phase, the total loss function of TopoSeg is

$$\mathcal{L}_{total} = \mathcal{L}_{pixel} + \lambda \mathcal{L}_{topo}, \tag{1}$$

where $\mathcal{L}_{pixel}$ denotes pixel-wise loss with respect to the output of both branches, $\mathcal{L}_{topo}$ denotes the proposed topology-aware in TAM with respect to the output at three-class branch and λ denotes the weight of topology-aware loss.

The design of this total loss leverages the strengths of both types of losses. Initially, the model is trained with only the pixel-wise loss for a certain number of iterations ($\lambda = 0$) to obtain accurate pixel-level predictions. Subsequently, the topology loss is introduced ($\lambda > 0$) to further enhance the model's ability to capture rational object structure.

In the inference phase, the final predicted instance segmentation map is obtained by utilizing boundaries predicted in the three-class branch to separate the semantic result of the semantic branch.

3.2. Topology Encoding

As previously noted, pixel-wise loss is insensitive to the differences in a few pixels at the boundary between overlapping nuclei, as distinct examples depicted in Fig. 3, despite significant structural differences. To quantitatively represent them, we present a topology encoding approach whose general representation is as follows.

For image with K classes, consider pixels of class i as set S_i. The zero-dimensional topology structure refers to the connection components formed by the connected pixels in S_i, and the one-dimensional structure refers to the holes. The topology of S_i can be considered as a family of subsets $\{S^d_{ij} | d = 0, 1; j = 1, 2, ..., b_d\}$ formed by topological structures of different dimensions, where d represents topology dimension and b_d represents the number of subsets of dimension d. Then a subset S^0_{ij} represents a connected component and S^1_{ij} represents a hole, Fig. 3 (Bottom).

To quantify the topological characteristics of S_i, we introduce the Betti number from algebraic topology. For a 2D image, the Betti numbers can be expressed as $B = (b_0, b_1)$, where b_0 denotes the number of connected components and b_1 denotes the number of 2D holes. Given a K-class image, the topology encoding can be formulated as

$$\mathbf{B} = \{B^d_i | i = 1, 2, ..., K; d = 0, 1\}, \tag{2}$$

where B^d_i represents Betti number of subset S^d_i.

In the nuclear instance segmentation task, the topology encoding of instance maps are mapped to three-class maps with three categories (inside, boundary and their union):

$$\mathbf{B} = \begin{bmatrix} B_{in} \\ B_{bo} \\ B_{un} \end{bmatrix} = \begin{bmatrix} N_{in} & 0 \\ N_{bo} & N_{in} \\ N_{bo} & 0 \end{bmatrix}, \tag{3}$$

where B_{in}, B_{bo} and B_{un} respectively represent the Betti numbers of the corresponding class. The specific values of B^d_i are defined based on the number of instances N_{in} and non-connected boundaries N_{bo} in the image.

Unlike the fixed topology prior of semantic given in [10] and [1], our topology encoding represents dynamically changing instance information for each image through three channels. As an example, the corresponding topology encoding of images in Fig. 3 (Top) is:

$$\begin{aligned} \mathbf{B}(GT) &= \begin{bmatrix} 2 & 0 \\ 1 & 2 \\ 1 & 0 \end{bmatrix}, \mathbf{B}(P1) = \begin{bmatrix} 2 & 0 \\ 1 & 2 \\ 1 & 0 \end{bmatrix}, \\ \mathbf{B}(P2) &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}, \mathbf{B}(P3) = \begin{bmatrix} 2 & 0 \\ 2 & 2 \\ 2 & 0 \end{bmatrix}. \end{aligned} \tag{4}$$

3.3. Topology-Aware Module

3.3.1 Persistence Encoding

Consider a probability map of class i as a topology set S_i with topology subsets $\{S^d_{ij} | d = 0, 1; j = 1, 2, ..., J_d\}$, where J_d represents the total number of subsets of dimension d. Here we use persistent homology to extract meaningful topology structures which are robust to noise and small perturbations in the probability map.

The first step of persistent homology is to construct a sequence of combinations $\{C_n\}$ of subsets, where each combination is obtained by adding more and more elements from S_i, which can be obtained by thresholding the probability map at a sequence of decreasing threshold values from 1 to 0. During this process, some subsets disappear with the gradual increase of pixels, while others appear. In Fig. 4, subsets of different dimensions are marked with letters of different colors. The subset a and b both appear in C_1, and a survives since then while b disappears in C_2 because the added pixels connect b and a into a common subset.

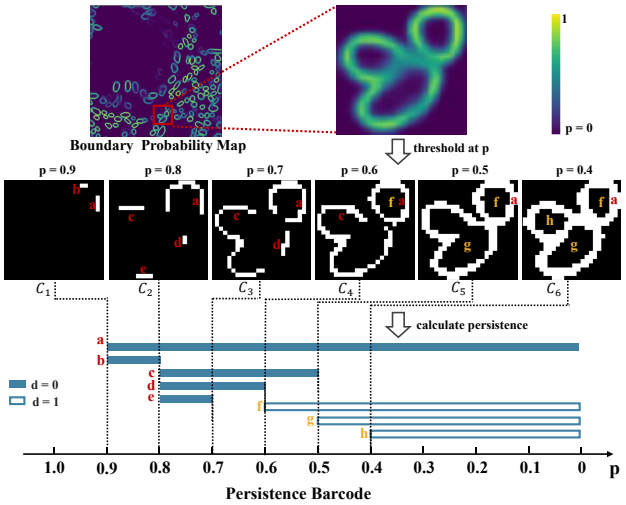

Figure 4. **The binary images obtained by taking different threshold p of the boundary probability map and corresponding persistence barcode for topology structures.** The red letters indicate connected components ($d = 0$), which is signified by the blue solid bars in the persistence barcode map. The golden letters indicate holes ($d = 1$), which are signified by the blue open bars.

Then the persistence of S_{ij}^d is measured by how long it survives in the sequence of combinations, which can be expressed through the persistence barcode whose length represents the corresponding persistence, Fig. 4. The longer the persistence barcode, the more robust the subset. Thereby, the persistence P_{ij}^d of S_{ij}^d can be calculated as the difference between the threshold values at which S_{ij}^d first appears ($p_{birth}(S_{ij}^d)$) and when it disappears ($p_{death}(S_{ij}^d)$):

$$P_{ij}^d = p_{birth}(S_{ij}^d) - p_{death}(S_{ij}^d). \tag{5}$$

Given prediction probabilities with K classes of a 2D image, the persistence encoding P is

$$\mathbf{P} = \{P_{ij}^d | i = 1, 2, ..., K; d = 0, 1; j = 1, 2, ..., J_d\}. \tag{6}$$

Note that since ground truth can be regarded as a probability map containing only binary values of 0 and 1, the persistence encoding of it is

$$\mathbf{Y} = \{Y_{ij}^d | i = 1, 2, ..., K; d = 0, 1; j = 1, 2, ..., B_i^d\}, \tag{7}$$

where the value of Y_{ij}^d is 1. As predicted probability maps typically contain more topological structures than those in ground truth, we assume that $J_d > B_i^d$.

3.3.2 Topology-Aware Loss

Based on persistence encoding, a better prediction should meet the condition that the subsets of C_n consistent with ground truth have greater persistence. We match the top B_i^d persistence barcode P_{ij}^d with GT and regard the remaining codes as unmatched. Our purpose is to reduce the error between the persistence encoding of prediction and ground truth, so as to make the topology correctness of prediction more robust. Therefore, the difference between them can be calculated in two parts:

$$\begin{aligned} M_i &= \sum_{d\in\{0,1\}} \sum_{j=1}^{B_i^d} (Y_{ij}^d - P_{ij}^d), \\ U_i &= \sum_{d\in\{0,1\}} \sum_{j=B_i^d+1}^{J_d} (P_{ij}^d), \end{aligned} \tag{8}$$

where M_i measures the length difference between the matched barcode and GT barcode for the ith class, and U_i measures the length of the unmatched barcode.

Given prediction probabilities with K classes of a 2D image, the topology-aware loss $\mathcal{L}_{topo}$ is

$$\mathcal{L}_{topo} = \frac{1}{K} \sum_{i=1}^{K} (M_i + U_i). \tag{9}$$

Minimizing this loss is equivalent to maximizing the persistence of topology features that are consistent with ground truth, and minimizing the persistence of undesired features. This provides a structure perspective to optimize the model, which overcomes the limitation of pixel-wise loss and achieves superior nucleus segmentation performance.

3.4. Adaptive Topology-Aware Selection (ATS)

ATS aims at efficiently optimizing images with topology-important regions by providing image- and patch-level selection for loss calculation, Fig. 5.

At the image level, we calculate the topology error between prediction and GT, then selects and stores images with an adaptive threshold, Fig. 5 (Left). If the topology error exceeds the threshold, it indicates that the overall structure of the image is significantly different from ground truth, then the image is preferentially selected and stored for further learning. On the contrary, if the topology error does not exceed the threshold, it will be dropped out in the next iteration. The adaptive threshold is determined based on the current number of iterations and topology error rate. The more iterations, the higher the adaptive threshold.

At the patch level, we split the pipeline into three main parts, Fig. 5 (Right). First, we divide the original image into hierarchical patches of varying scales. Then the topology error between predicted patches and corresponding ground truth is calculated. Moreover, a sorted list of patches is built based on these. According to their topology errors, an adaptive screen method is adopted. This allows the calculation of topology-aware loss on selected patches, which drives the model to focus on regions with more topology errors.

In more detail, the Betti error is calculated by comparing the difference in the number of disjoint pixel sets of each

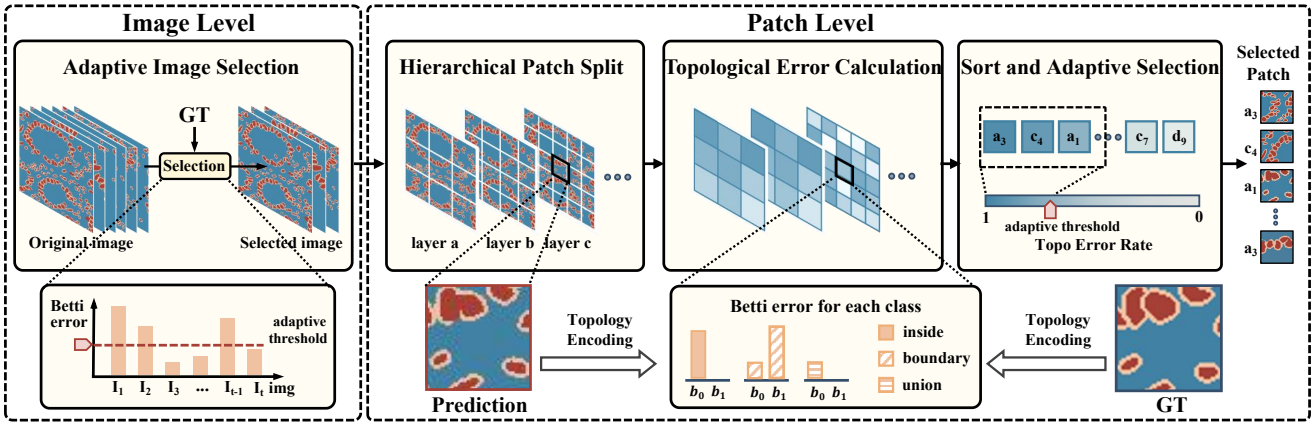

Figure 5. **The pipeline of adaptive topology-aware selection module (ATS)**. In the topological Error Calculation and Sort and Adaptive Selection, the darker the square, the larger the Betti error.

class between the prediction and GT, Fig. 5 (Bottom right). Note that the Betti error calculation used in the selection at the image and patch levels is the same. The difference is that one uses the whole image and the other uses the patch.

4. Experiment

4.1. Experimental Settings

Datasets. I) MoNuSeg [17] contains 30 H&E-stained pathological images of size 1000×1000 from seven organs, including a total of 21,623 nuclei labeled without distinction between various categories. According to [17], we divide 30 images into three sets: 12 for training, 4 for validation, and 14 for testing (combining the same organ and different organ test sets). II) CPM17 [27] contains 64 H&E stained histopathology images with 7,570 annotated nuclear boundaries. It is from the MICCAI 2017 Digital Pathology Challenge [27] and images on two different scales: 500×500 and 600×600. This is also a dataset for single-class segmentation task. III) PBNulei is a nuclear segmentation dataset from Peking University Affiliated Hospital, marked by professional pathologists. It contains 106 H&E stained histopathology images of 512×512 sizes, with a total of 36,659 nuclei, 78 images for training and 28 for testing. The dataset is divided into five categories, namely carcinoma, epithelial, inflammatory, mesenchymal, and myoepithelium.

Evaluation Metrics. We use three instance-level evaluation metrics to measure the instance segmentation performance of the comparison models, which are: mean Panoptic Quality (PQ), the Aggregated Jaccard Index (AJI), and average Hausdorff distance (Haus).

Implementation details. The proposed TopoSeg is trained and tested using the open-source software library Pytorch 1.8.0 on NVIDIA Tesla V100 with CUDA 10.1.

The MoNuSeg dataset has a relatively large original image size and the CPM17 dataset has an inconsistent original image size, so it is divided into multiple patches, thus allowing the network to be uniform and fully trained. As a result, the MoNuSeg dataset generates 192 training patches, the

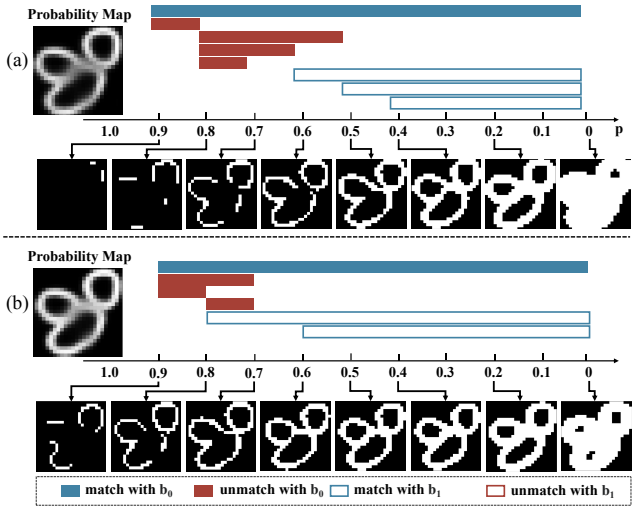

Figure 6. **Qualitative comparison before and after adding the topology-aware loss**. (a) the probability map obtained by the model just using pixel-wise loss; (b) the probability map obtained by the model using pixel-wise loss and topology-aware loss.

Dataset	Backbone	Loss		PQ ↑	AJI ↑	Haus ↓
		pixel	topo			
MoNuSeg	U-Net	✓		0.581	0.579	12.80
	U-Net	✓	✓	**0.597**	**0.619**	**9.647**
CPM	U-Net	✓		0.625	0.666	7.740
	U-Net	✓	✓	**0.662**	**0.718**	**5.961**
PBNulei	U-Net	✓		**0.426**	0.386	2.770
	U-Net	✓	✓	0.425	**0.416**	**1.953**

Table 1. Comparison of instance segmentation results obtained using only pixel-wise loss and adding topology-aware loss based on pixel-wise loss in three datasets.

CPM17 dataset generates 512 training patches. In the training phase, we apply a series of data augmentation methods, which include random flip, random rotation, random color, random crop, and Gaussian blur. The training epoch is set as 300, and the initial learning rate is set as 10^{-3}.

4.2. Ablation studies

We conducted ablation studies using three datasets, MoNuSeg, CPM17, and PBNulei, to validate the effectiveness of the proposed TopoSeg. We selected U-Net as the baseline segmentation method, and we improved the U-Net by adding two output branches: a semantic prediction branch and a three-class prediction branch.

Topology-Aware Loss. When only pixel-wise losses are present, the model's performance on the MoNuSeg dataset is 0.581 PQ, 0.579 AJI, and 12.8 Haus, Table 1. After incorporating topological awareness losses, the model's PQ, AJI, and Haus on the MoNuSeg dataset are 0.597, 0.619, and 9.647, respectively, which have increased by 1.7%, 3.9%, and 3.153 compared to the original results.

Compared to the performance obtained by using only pixel-wise loss, the model trained with the addition of pixel-wise and topological loss has increased by 3.7% PQ, 5.2% AJI, and 1.779 Haus respectively on the CPM17 dataset and by 3.0% AJI, and 0.817 Haus respectively on the PBNulei dataset compared to the original results.

Fig. 6 makes a qualitative comparison between the boundary probability map obtained before and after adding the topology-aware loss and its corresponding persistence barcode. It shows that the corresponding structure obtained after adding topology-aware loss is more persistent. As an example, when p=0.5, three instances are correctly segmented, but when only pixel-wise loss is used, one instance can not be predicted. This indicates that adding topology-aware loss on the basis of pixel-wise loss helps the model learn object structure information, thus more accurately segmenting nuclear instances.

Adaptive Image Selection. We set the threshold of 25%, 50%, 75% and the adaptive threshold for the comparisons. According to the threshold, select the image with Betti error greater than the threshold as the new training dataset, and ensure that the number of training data obtained by different thresholds is the same.

In Table 2, we compare the performance before and after image selection. The comparison results show that the segmentation results obtained after image selection are better than those without image selection. This proves that image selection promotes the model to focus on learning images with more structural prediction errors and improves the instance segmentation performance of the model.

Compared to training with data selected with a threshold of 25% or 75%, training with data selected with a threshold of 50% yielded more competitive model performance,

Datasets	Backbone	S_{img}	PQ ↑	AJI ↑	Haus ↓
MoNuSeg	U-Net	100%	0.597	0.619	9.647
		75%	0.604	0.626	9.358
		50%	0.606	0.628	9.328
		25%	0.602	0.625	9.452
		Adaptive	**0.610**	**0.632**	**9.113**
CPM	U-Net	100%	0.662	0.718	5.961
		75%	0.665	0.723	5.835
		50%	0.670	0.724	5.887
		25%	0.672	0.723	5.847
		Adaptive	**0.679**	**0.731**	**5.805**
PBNulei	U-Net	100%	0.425	0.416	1.953
		75%	0.427	0.419	2.014
		50%	0.419	0.414	2.028
		25%	0.429	0.418	2.009
		Adaptive	**0.446**	**0.435**	**1.931**

Table 2. **Comparison results of the models obtained by selecting images based on different thresholds.** Here, 100% represents the use of raw data without any image selection operation. 25%, 50%, and 75% represent selecting images with threshold values greater than 25%, 50%, and 75%, respectively, as new input data. Adaptive represents the proposed adaptive image selection method.

Datasets	Backbone	S_{patch}	PQ ↑	AJI ↑	Haus ↓
MoNuSeg	U-Net	Whole	0.597	0.619	9.647
		Non-hierc	0.609	0.628	10.063
		Hierarchical	0.613	0.639	8.842
		Adaptive	**0.616**	**0.643**	**8.787**
CPM	U-Net	Whole	0.662	0.718	5.961
		Non-hierc	0.675	0.729	6.108
		Hierarchical	0.662	0.733	**5.510**
		Adaptive	**0.678**	**0.741**	5.605
PBNulei	U-Net	Whole	0.425	0.416	**1.953**
		Non-hierc	0.498	0.479	2.320
		Hierarchical	0.514	0.499	2.011
		Adaptive	**0.520**	**0.504**	2.185

Table 3. **Ablation studies on different patch selection methods.** S_{patch} denotes the method of dividing the image, "Whole" denotes the whole image for calculation, "Non-hierc" means non-hierarchical and refers to single size patch division, "Hierarchical" refers to multi-scale patch division, and "Adaptive" refers to the adaptive approach of dividing the patch.

indicating that overly high or low thresholds can limit the learning of the model. Therefore, selecting a single threshold for image selection has limitations.

Compared to the three fixed threshold image filtering methods, our method has the ability to flexibly select input data. As shown in Table 2, the adaptive image selection approach achieved superior segmentation performance. In the MoNuSeg dataset, the performance metrics PQ, AJI,

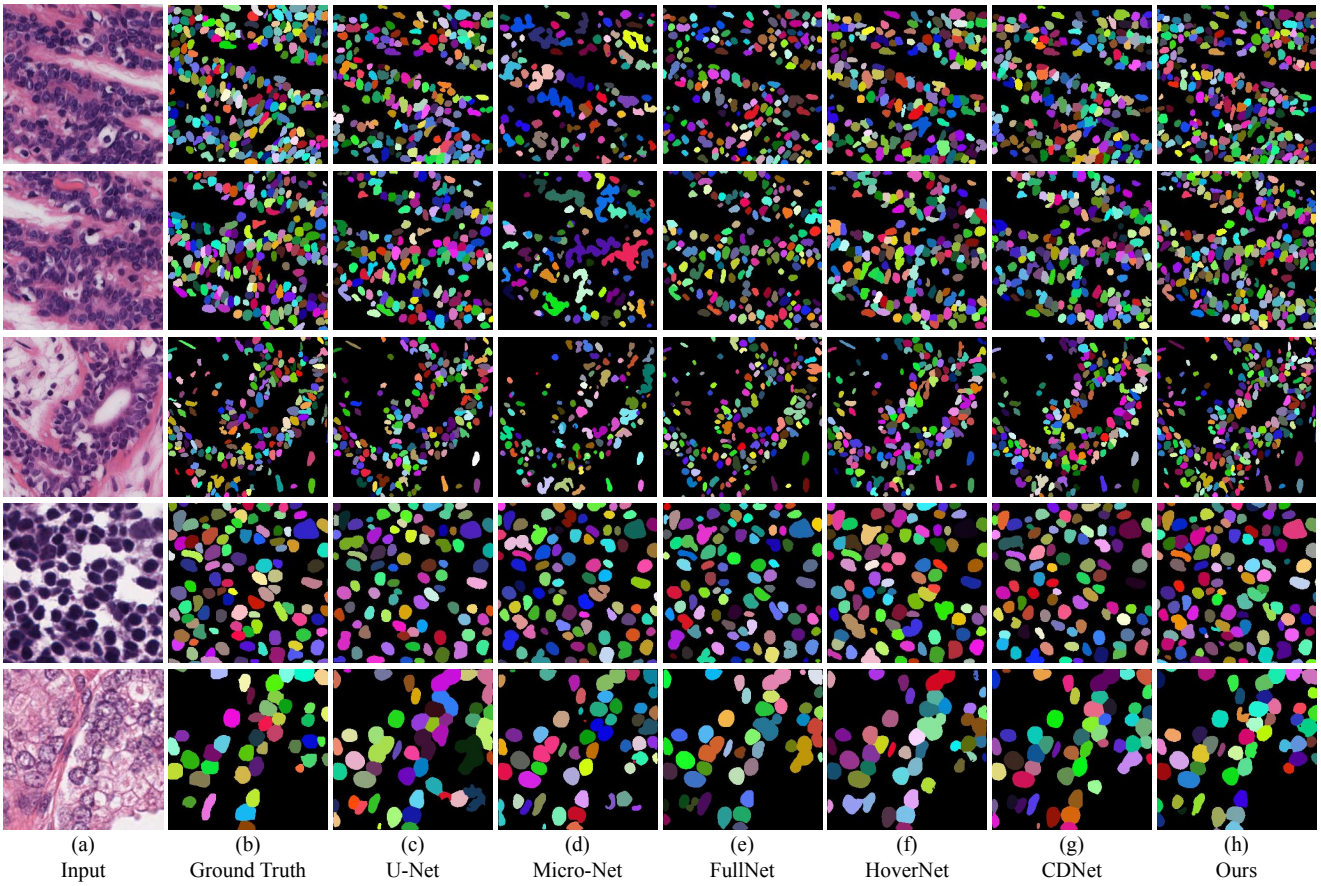

Figure 7. **Visualization of instance segmentation results with different methods.** (a) Input image; (b) Ground Truth; (c) U-Net [25]; (d) Micro-Net [24]; (e) FullNet [23]; (f) Hover-Net [9]; (g) CDNet [11], (h) Our TopoSeg. Different colors indicate different instances in the images. The white rectangles are drawn for a clear comparison.

and Haus improved by 1.2%, 1.3%, and 0.534, respectively, compared to the original results. Similarly, compared with the original results, the performance of the CPM17 dataset is improved by 1.7% PQ, 1.3% AJI and 0.16 Haus respectively, and the performance of the PBNulei dataset is improved by 2% PQ, 1.9% AJI and 0.022 Haus, respectively.

This reflects that, compared with the three fixed threshold image selection methods, our adaptive image selection approach has the ability to select input data more flexibly, thus promoting the model to obtain more correct instance segmentation results.

Adaptive Hierarchical Patch Selection. As shown in Table 3, we compare the performance of the topology loss calculation after patch selection with that of the whole image. First of all, we verify the effect of patch selection on segmentation performance. Comparing the first row (Whole) and the second row (Non-hierc) in each dataset in Table 3, it is concluded that the segmentation performance of topology loss calculation after selecting patch is better than that of using the whole image on both PQ and AJI. This shows that by selecting the appropriate patch, the model pays more attention to the regions that need structural optimization, so as to improve the segmentation performance.

Then, we verified the effect of using hierarchical patches with different scales for calculating topological loss in instance segmentation. We compared the performance of the third row (Hierarchical) and the second row (Non-hierc) within each dataset in Table 3 and found that hierarchical patches outperformed single-layer patches. This also shows that selecting the appropriate region that needs structural optimization helps the model better segment instances.

Table 3 shows that compared with the topology loss calculated by the whole image, the segmentation performance obtained by adaptive hierarchical selection of patch is significantly improved. In the MoNuSeg dataset, PQ increased by 1.8%, AJI increased by 2.4%, Haus increased by 0.86. In the CPM dataset, PQ increased by 1.6%, AJI increased by 2.3%, Haus increased by 0.356. In the PBNulei dataset, PQ increased by 9.5%, AJI increased by 8.8%.

Method	MoNuSeg		CPM17		PBNuclei	
	PQ ↑	AJI ↑	PQ ↑	AJI ↑	PQ ↑	AJI ↑
U-Net [25]	0.581	0.579	0.625	0.666	0.355	0.312
Mask-RCNN [13]	0.509	0.546	0.674	0.684	-	-
DCAN [2]	0.492	0.525	0.545	0.561	0.243	0.249
DIST [21]	0.443	0.559	0.504	0.616	0.325	0.301
Micro-Net [24]	0.519	0.56	0.661	0.668	0.306	0.287
Full-Net [23]	-	0.604	-	0.661	0.385	0.378
PFF-Net [18]	0.587	0.611	-	-	-	-
Hover-Net [9]	0.597	0.618	0.697	0.705	0.416	0.363
CDNet [11]	-	0.637	-	0.733	0.413	0.381
Ours	**0.625**	**0.643**	**0.705**	**0.756**	**0.528**	**0.509**

Table 4. Performance comparisons on three nuclei datasets.

4.3. Performance Comparison

In Table 4, the proposed TopoSeg is compared with the state-of-the-art nuclear segmentation model on MoNuSeg and CPM17. Our TopoSeg achieves 0.625 PQ and 0.643 AJI on the MoNuSeg dataset and 0.705 PQ and 0.756 AJI on CPM17 dataset, respectively. The TopoSeg achieves competitive results for nuclei segmentation task.

We further carry out a qualitative visualization analysis of our method and different existing methods. Compared to all the other methods in Fig. 7, our TopoSeg achieves better instance segmentation results for nuclear instance segmentation, especially for the correct structure segmentation of dense or overlapping regions. Both quantitative and qualitative results demonstrate that our TopoSeg is competitive and able to segment dense nuclei instances.

5. Conclusion

This paper proposes a novel topology-aware segmentation approach, termed TopoSeg, for nuclear instance segmentation tasks to address the challenge of touching and overlapping nuclear segmentation. We develop a TAM, which contains a persistence encoding method that accounts for topology information in predicted probability maps, and a topology-aware loss that is flexible and suitable for topology optimization. We further improve it by applying an ATS, which enforces the model to focus on regions with high topology errors. The experimental results in three nuclei datasets show that TopoSeg has achieved superior performance. As a general method of topology awareness, TopoSeg provides fresh insight into the instance segmentation problem of small and dense objects.

Acknowledgments This work was supported in part by the National Key R&D Program of China (No. 2022ZD0118201), Natural Science Foundation of China (No. 61972217, 32071459, 62176249, 62006133, 62271465).

References

[1] Nick Byrne, James R Clough, Israel Valverde, Giovanni Montana, and Andrew P King. A persistent homology-based topological loss for cnn-based multiclass segmentation of cmr. *IEEE Transactions on Medical Imaging*, 42(1):3–14, 2022. 2, 3, 4

[2] Hao Chen, Xiaojuan Qi, Lequan Yu, Qi Dou, Jing Qin, and Pheng-Ann Heng. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical Image Analysis*, 36:135–146, 2017. 2, 9

[3] Shengcong Chen, Changxing Ding, Minfeng Liu, Jun Cheng, and Dacheng Tao. Cpp-net: Context-aware polygon proposal network for nucleus segmentation. *IEEE Transactions on Image Processing*, 32:980–994, 2023. 1, 2

[4] James R Clough, Nicholas Byrne, Ilkay Oksuz, Veronika A Zimmer, Julia A Schnabel, and Andrew P King. A topological loss function for deep-learning based image segmentation using persistent homology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12):8766–8778, 2020. 3

[5] James R Clough, Ilkay Oksuz, Nicholas Byrne, Julia A Schnabel, and Andrew P King. Explicit topological priors for deep-learning based image segmentation using persistent homology. In *International Conference on Information Processing in Medical Imaging*, pages 16–28, 2019. 3

[6] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, pages 454–463, 2000. 2

[7] Robin Forman. Morse theory for cell complexes. *Advances in Mathematics*, 134(1):90–145, 1998. 3

[8] Robin Forman. A user's guide to discrete morse theory. *Séminaire Lotharingien de Combinatoire [electronic only]*, 48:1–35, 2002. 3

[9] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis*, 58:1–15, 2019. 1, 2, 8, 9

[10] Saumya Gupta, Xiaoling Hu, James Kaan, Michael Jin, Mutshipay Mpoy, Katherine Chung, Gagandeep Singh, Mary Saltz, Tahsin Kurc, Joel Saltz, et al. Learning topological interactions for multi-class medical image segmentation. In *Proceedings of the European Conference on Computer Vision*, pages 701–718, 2022. 4

[11] Hongliang He, Zhongyi Huang, Yao Ding, Guoli Song, Lin Wang, Qian Ren, Pengxu Wei, Zhiqiang Gao, and Jie Chen. Cdnet: Centripetal direction network for nuclear instance segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4026–4035, 2021. 2, 8, 9

[12] Hongliang He, Chi Zhang, Jie Chen, Ruizhe Geng, Luyang Chen, Yongsheng Liang, Yanchang Lu, Jihua Wu, and Yongjie Xu. A hybrid-attention nested unet for nuclear segmentation in histopathological images. *Frontiers in Molecular Biosciences*, 8:1–9, 2021. 1, 2

[13] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2961–2969, 2017. 9

[14] Xiaoling Hu, Fuxin Li, Dimitris Samaras, and Chao Chen. Topology-preserving deep image segmentation. *Advances in Neural Information Processing Systems*, 32:5658–5669, 2019. 2, 3

[15] Xiaoling Hu, Dimitris Samaras, and Chao Chen. Learning probabilistic topological representations using discrete morse theory. *arXiv preprint arXiv:2206.01742*, 2022. 3

[16] Xiaoling Hu, Yusu Wang, Li Fuxin, Dimitris Samaras, and Chao Chen. Topology-aware segmentation using discrete morse theory. In *Proceedings of the International Conference on Learning Representation*, pages 1–19, 2021. 2, 3

[17] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging*, 36(7):1550–1560, 2017. 6

[18] Dongnan Liu, Donghao Zhang, Yang Song, Heng Huang, and Weidong Cai. Panoptic feature fusion net: A novel instance segmentation paradigm for biomedical and biological images. *IEEE Transactions on Image Processing*, 30:2045–2059, 2021. 1, 9

[19] Xiaoming Liu, Zhengsheng Guo, Jun Cao, and Jinshan Tang. Mdc-net: a new convolutional neural network for nucleus segmentation in histopathology images with distance maps and contour information. *Computers in Biology and Medicine*, 135:1–12, 2021. 1, 2

[20] Agata Mosinska, Pablo Marquez-Neila, Mateusz Koziński, and Pascal Fua. Beyond the pixel-wise loss for topology-aware delineation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3136–3145, 2018. 2

[21] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Transactions on Medical Imaging*, 38(2):448–459, 2018. 1, 2, 9

[22] H. Pinckaers, W. Bulten, J. van der Laak, and G. Litjens. Detection of prostate cancer in whole-slide images through end-to-end training with image-level labels. *IEEE Transactions on Medical Imaging*, 40(7):1817–1826, 2021. 1

[23] Hui Qu, Zhennan Yan, Gregory M Riedlinger, Subhajyoti De, and Dimitris N Metaxas. Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 378–386, 2019. 1, 2, 8, 9

[24] Shan E Ahmed Raza, Linda Cheung, Muhammad Shaban, Simon Graham, David Epstein, Stella Pelengaris, Michael Khan, and Nasir M Rajpoot. Micro-net: A unified model for segmentation of various objects in microscopy images. *Medical Image Analysis*, 52:160–173, 2019. 2, 8, 9

[25] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241, 2015. 1, 2, 8, 9

[26] Suprosanna Shit, Johannes C Paetzold, Anjany Sekuboyina, Ivan Ezhov, Alexander Unger, Andrey Zhylka, Josien PW Pluim, Ulrich Bauer, and Bjoern H Menze. cldice-a novel topology-preserving loss function for tubular structure segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16560–16569, 2021. 2

[27] Quoc Dang Vu, Simon Graham, Tahsin Kurc, Minh Nguyen Nhat To, Muhammad Shaban, Talha Qaiser, Navid Alemi Koohbanani, Syed Ali Khurram, Jayashree Kalpathy-Cramer, Tianhao Zhao, et al. Methods for segmentation and classification of digital microscopy tissue images. *Frontiers in Bioengineering and Biotechnology*, 53:1–15, 2019. 6

[28] Yanning Zhou, Omer Fahri Onder, Qi Dou, Efstratios Tsougenis, Hao Chen, and Pheng-Ann Heng. Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation. In *International Conference on Information Processing in Medical Imaging*, pages 682–693, 2019. 2