# Unleashing the Power of Prompt-driven Nucleus Instance Segmentation

Zhongyi Shui[1,2], Yunlong Zhang[1,2], Kai Yao[3], Chenglu Zhu[2], Sunyi Zheng[2], Jingxiong Li[1,2], Honglin Li[1,2], Yuxuan Sun[1,2], Ruizhe Guo[1,2], and Lin Yang[2,†]

[1] College of Computer Science and Technology, Zhejiang University
[2] School of Engineering, Westlake University
[3] University of Liverpool
{shuizhongyi,yanglin}@westlake.edu.cn
[†]Corresponding Author

**Abstract.** Nucleus instance segmentation in histology images is crucial for a broad spectrum of clinical applications. Current dominant algorithms rely on regression of nuclear proxy maps. Distinguishing nucleus instances from the estimated maps requires carefully curated post-processing, which is error-prone and parameter-sensitive. Recently, the Segment Anything Model (SAM) has earned huge attention in medical image segmentation, owing to its impressive generalization ability and promptable property. Nevertheless, its potential on nucleus instance segmentation remains largely underexplored. In this paper, we present a novel prompt-driven framework that consists of a nucleus prompter and SAM for automatic nucleus instance segmentation. Specifically, the prompter is developed to generate a unique point prompt for each nucleus, while SAM is fine-tuned to produce its corresponding mask. Furthermore, we propose to integrate adjacent nuclei as negative prompts to enhance model's capability to identify overlapping nuclei. Without complicated post-processing, our proposed method sets a new state-of-the-art performance on three challenging benchmarks. Code available at https://github.com/windygoo/PromptNucSeg.

**Keywords:** Pathology image analysis · Nucleus instance segmentation

## 1 Introduction

Cancer is one of the leading causes of death worldwide. Over the past decades, substantial endeavors have been made to detect cancers from histology images with the aim of improving survival rates through early screening. Identification of nuclear components in the histology landscape is often the first step toward a detailed analysis of histology images. Quantitative characterizations of nuclear morphology and structure play a pivotal role in cancer diagnosis, treatment planning, and survival analysis, which have been verified by a wide range of studies, see for example [1]. However, large-scale analysis on the cell level is extremely labor-intensive and time-consuming since a whole slide image (WSI) typically

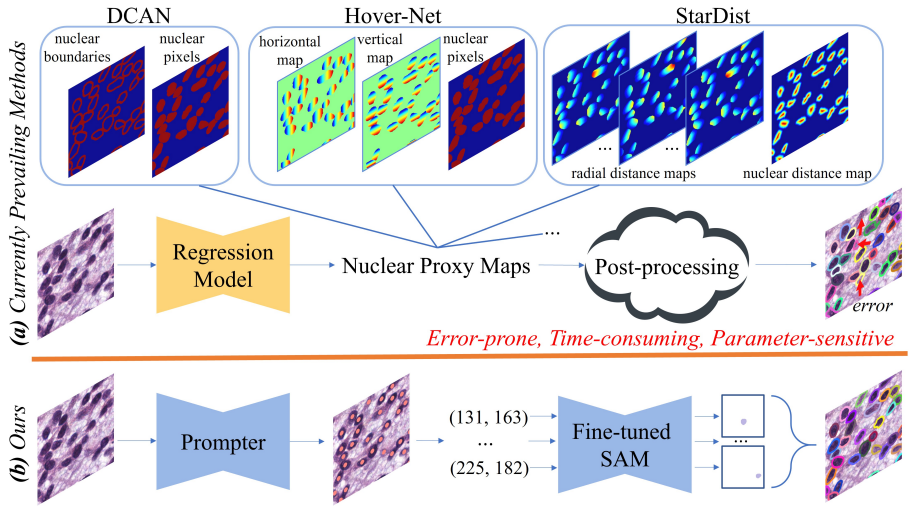


**Fig. 1:** Pipeline comparison with currently prevailing nucleus instance segmentation algorithms.

contains tens of thousands of nuclei of various types. Moreover, such subjective interpretations have been demonstrated to suffer from large inter-and intra-observer variability [11]. Consequently, there is a compelling pursuit of precise automatic algorithms for nucleus instance segmentation to aid in histopathologic cancer diagnosis. Nonetheless, the blurred cell contours, overlapping cell clusters, and variances in nuclei staining, shape and size, pose substantial challenges for the developers.

Recent years have witnessed significant advancements in the filed of nucleus instance segmentation owing to the impressive performances brought by various methods based on regression of nuclear proxy maps [2–4, 10, 16, 27, 31, 41, 41] (see Fig. 1 (a)). Regrettably, these methods necessitate carefully crafted post-processing to derive nuclear instances from the estimated maps. This step demands meticulous hyper-parameter tuning and is vulnerable to noise [37].

Recently, the segment anything model (SAM) has emerged as a generic segmentation network for various image types, whose impressive generalization ability and versatility can be attributed to its structural design and the strong representation learned from 11M images annotated with 1B masks [17]. Several studies have been undertaken to investigate the zero-shot performance of SAM on nucleus segmentation [6] or transfer its well-learned representation to boost the segmentation accuracy [14,36]. Specifically, [14] reuses SAM's well-trained image encoder to construct a more powerful regression model and integrates it into the aforementioned nucleus instance segmentation workflow. Despite the promising results, we argue that this approach does not fully exploit the knowledge encapsulated in the integrated architecture of SAM. Conversely, [36] maintains the

philosophy of SAM thoroughly. They fine-tune the entire SAM in a one-prompt-all-nuclei recipe for nucleus semantic segmentation. Nevertheless, this method expects users to supply precise prompts, which is impractical since crafting such prompts requires extensive medical expertise. Moreover, it falls short in providing nucleus instance information.

In this work, we propose to fine-tune SAM in a one-prompt-one-nucleus regime to fully unleash its potential for nucleus instance segmentation. To eliminate the need for crafted prompts during inference, we develop a prompter that automatically generates nuclei prompts by refining and classifying pre-defined anchor points on an input image. Specially, we incorporate an auxiliary task of nuclear region segmentation into prompter learning. This integration guides the model's attention towards foreground areas, thereby improving the quality of generated prompts. During the inference stage, the predicted nuclear region mask is further utilized to filter out false positive prompts. The consolidation of the prompter and segmentor (i.e., the fine-tuned SAM) establishes a novel solution for automatic nucleus instance segmentation. Given their linkage through nuclei prompts, we designate our approach as PromptNucSeg, and its pipeline is depicted in Fig. 1 (b). Compared to the currently prevailing methods, our approach does not require complex post-processing. Moreover, we devise a trick that treats adjacent nuclei as negative prompts to improve the model's segmentation of overlapping nuclei.

Our contributions can be summarized as follows:

- We propose PromptNucSeg, which provides a new perspective for nucleus instance segmentation.
- We develop a prompter for automatic nuclei prompts generation and design a simple auxiliary task to boost its performance.
- We propose to use adjacent nuclei as negative prompts to promote segmentation of overlapping nuclei.
- Extensive experiments on three challenging benchmarks demonstrate the advantages of PromptNucSeg over the state-of-the-art counterparts.

## 2 Related Work

### 2.1 Utilization of SAM for Medical Image Segmentation

Segment Anything Model (SAM) [17] is the first groundbreaking model for universal image segmentation. It has achieved impressive results on a wide range of natural image tasks. Nevertheless, due to the dramatic domain gap between natural and medical images, SAM's performance significantly declines when applied for medical image segmentation [15, 24]. To bridge this gap, many studies opt to fine-tine SAM with meticulously curated medical data [5, 19, 22, 34, 35, 39]. These works mainly focus on the segmentation of anatomical structures and organs in computed tomography, magnetic resonance and ultrasound images.

In terms of histology images, [6] assesses SAM's performance for tumor, non-tumor tissue and nucleus segmentation. The results suggest that the vanilla

SAM achieves remarkable segmentation performance for large connected tissue objects, however, it does not consistently achieve satisfactory results for dense nucleus instance segmentation. To tackle this issue, SPPNet [36] fine-tunes a distilled lightweight SAM [38] in a one-prompt-all-nuclei manner for nucleus semantic segmentation. Despite the improved outcomes, this method depends on manual prompts and fails to furnish nucleus instance information. Although a recent study SAC [26] obviates the necessity for manual prompts, it is still confined to nucleus semantic segmentation. CellViT [14] builds a vision transformer-based U-Net-shaped model, employing SAM's pre-trained image encoder as its backbone to better fit the nuclear proxy maps. We argue that this approach underutilizes the knowledge embedded in SAM's integrated architecture.

### 2.2 Nucleus Instance Segmentation

Current methods for nucleus instance segmentation can be divided into two categories: top-down and bottom-up.

Top-down methods, such as Mask R-CNN [13], first predict nuclei bounding boxes from a global perspective, and then segment the nucleus instance within each box. Despite the great progress in natural image segmentation and the potential in dealing with overlapping nuclei, top-down methods have demonstrated deficiency on nucleus instance segmentation [10, 23, 37], attributed to two primary factors. First, on the data side, there are many severely overlapping nuclei in histology images. Consequently, a bounding-box proposal normally contains multiple nuclei with indistinct boundaries, making the network hard to optimize. Second, on the model side, top-down methods typically predict segmentation masks with a fixed resolution (e.g., 28×28 in Mask R-CNN). Subsequently, these masks undergo re-sampling to match the size of their corresponding bounding boxes. This re-sampling process might introduce quantization errors [37], posing challenges for accurately segmenting sinuous nuclear boundaries.

Bottom-up methods, initially regressing various types of nuclear proxy maps and then grouping pixels into individual instances through meticulous post-processing, have gained prominence in nucleus instance segmentation owing to their commendable accuracy. These approaches typically entail regressing a nucleus probability map, where the pixel values signify the presence of nuclei, along with some auxiliary maps facilitating the identification of nuclei instances. Specifically, DCAN [2], CIA-Net [41], TSFD-Net [16] and HARU-Net [3] predict the nuclear contour map. DIST [27] regresses the intra-nuclear distance map. HoVer-Net [10] predicts horizontal and vertical distances of nuclei pixels to their center of mass. StarDist [31] and its extension CPP-Net [4] predict distances from each foreground pixel to its associated instance boundary along a set of pre-defined directions. Under the premise of some above frameworks, other works [7, 14, 28, 40] put effort into constructing more favorable features or task-specific loss functions. Overall, while bottom-up methods have exhibited superior accuracy compared to top-down approaches, their accompanying post-processing requires tedious hyper-parameter tuning [37], which presents a hurdle to their practical application.

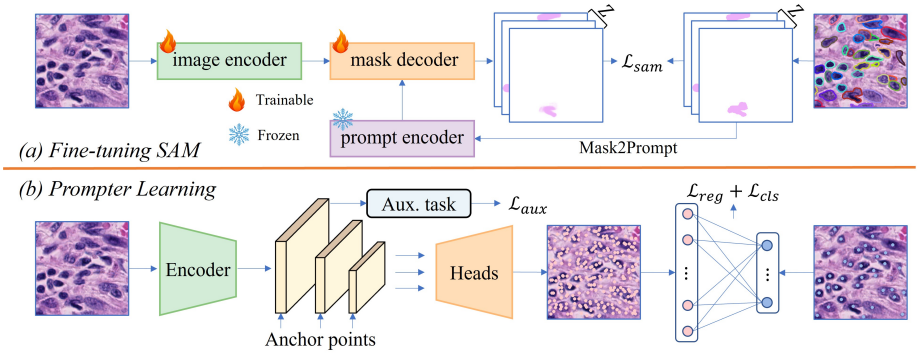


**Fig. 2:** (a) The fine-tuning process of SAM. Mask2Prompt signifies randomly sampling a positive point prompt from the foreground area of each nucleus mask. (b) The training procedure of the nucleus prompter. The integration of these two models enables automatic nucleus instance segmentation, as illustrated in Fig. 1 (b).

Essentially, our proposed PromptNucSeg belongs the top-down family. But inspired by the promptable property of SAM, we tackle this task from a new perspective. Instead of bounding boxes, we utilize center points to represent nuclei, which are easier to localize and can separate touching objects more precisely. In comparison with bottom-up methods, PromptNucSeg does not require intricate post-processing as the prompter generates point prompts for nuclei in a one-to-one relationship and the segmentor predicts the nuclei mask guided by each prompt individually.

## 3 Methodology

### 3.1 Preliminaries: SAM

SAM [17] consists of three sub-networks, *i.e.*, image encoder $\mathcal{F}$, prompt encoder $\mathcal{P}$ and mask decoder $\mathcal{M}$. The image encoder transforms an input image $I \in \mathbb{R}^{H\times W\times 3}$ into an image embedding. The prompt encoder maps diverse prompts (*e.g.*, a set of positive/negative points, a rough box or mask, free-form text, or combinations thereof) into a compact prompt embedding. Positive prompts indicate regions representing the region-of-interest (ROI) object, whereas negative prompts emphasize areas that should be suppressed as background. Given the image and prompt embedding as input, the mask decoder generates the mask for the ROI object in conjunction with a confidence score (*i.e.*, an estimated IoU).

### 3.2 Adapt SAM for Nucleus Instance Segmentation

Despite SAM's remarkable segmentation performance across numerous natural images, recent studies have highlighted its subpar performance on medical images

due to the significant domain gap [5, 15]. A specific observation worth noting is that the objects in SAM's pre-training data are primarily captured in natural scenes, displaying nicely delineated boundaries, while the boundaries of organs or nuclei in medical images are often ambiguous [5,36]. To enhance the capability of SAM for nucleus segmentation, we fine-tune it on nucleus instance segmentation datasets to incorporate essential domain-specific knowledge into the model.

The fine-tuning procedure is depicted in Fig. 2 (a). Specifically, for each image-label pair $(x, y)$ in a mini-batch, we randomly select $Z$ nucleus instances from the instance map $y$. Subsequently, a positive point prompt is randomly sampled from the foreground area of each instance. Taking the image $x$ and the point prompt $p_z$ as input, we fine-tune SAM to predict the mask of $z$-th nucleus instance.

$$\widetilde{\mathcal{O}}_z = \mathcal{M}\left(\mathcal{F}\left(x\right), \mathcal{P}\left(\{p_z\}\right), [\text{mask}], [\text{IoU}]\right) \tag{1}$$

where [mask] and [IoU] separately represent the learnable mask and IoU token pre-set in SAM's mask decoder. $\widetilde{\mathcal{O}}_z$ denotes the predicted mask of the $z$-th nucleus. We supervise the mask and IoU prediction with the same loss as SAM.

$$\mathcal{L}_{sam} = \omega \text{FL}\left(\widetilde{\mathcal{O}}_z, \mathcal{O}_z\right) + \text{DL}\left(\widetilde{\mathcal{O}}_z, \mathcal{O}_z\right) + \text{MSE}\left(\tilde{\nu}, \nu\right) \tag{2}$$

where FL, DL and MSE stand for focal loss [21], dice loss [25] and mean-square-error loss, respectively. $\mathcal{O}_z$ is the ground-truth mask of the $z$-th nucleus, $\tilde{\nu}$ and $\nu$ signify the estimated and actual IoU between $\widetilde{\mathcal{O}_z}$ and $\mathcal{O}_z$, respectively. $\omega$ is a weight term. In this work, we opt to freeze the prompt encoder while updating the image encoder and mask decoder via gradient descent.

### 3.3 Learn Prompter

Generating a unique point prompt for each nucleus is de facto a non-trivial problem. In this study, we choose the nuclear centroid as its prompt for simplicity. To achieve automatic prompt generation, we draw inspiration from [32] and develop a prompter to predict nuclear centroid coordinates and categories by refining and classifying a set of anchor points placed on an input image. In the following content, we denote the set of anchor points as $\mathcal{A} = \{a_i\}_{i=1}^{M}$ and the set of ground-truth points as $\mathcal{B} = \{b_i\}_{i=1}^{N}$, where $b_i$ is extracted from $y$ as the centroid of the $i$-th nucleus.

The prompter learning procedure is depicted in Fig. 2 (b). Specifically, we begin with placing anchor points on an input image $x$ with a step of $\lambda$ pixels. Then, an image encoder $\mathcal{F}'$ is employed to construct hierarchical feature maps $\{P_j\}_{j=2}^{L}$ from $x$, where the size of $P_j$ is $(H/2^j, W/2^j)$. Following this, we apply the bilinear interpolation method to extract multi-scale feature vectors $\{f_{i,j}\}_{j=2}^{L}$ for anchor point $a_i$ according to its normalized coordinates on the feature pyramid. Finally, we concatenate $\{f_{i,j}\}_{j=2}^{L}$ and fed it into two dedicated MLP heads for decoding offsets $\delta_i$ and logits $q_i \in \mathbb{R}^{C+1}$ with respect to $a_i$, where $C$ is the number of nuclear categories and the extra class is background.

Since the goal of prompter is to associate a unique point prompt for each nucleus, which anchor point in $\mathcal{A}$ should be chosen as the prompt is the key in prompter learning. In principal, for any nucleus centroid in $\mathcal{B}$, the anchor point with lower distance and higher categorical similarity with it is preferred to be chosen. Consequently, the association can be completed by computing the maximum-weight matching $\phi = \{(a_{\sigma(i)}, b_i)\}_{i=1}^{N}$ in a weighted bipartite graph $\mathcal{G} = (\mathcal{A}, \mathcal{B}, \mathcal{E})$, where the weight $w_{i,j}$ of edge connecting vertex $a_i$ and $b_j$ is defined as:

$$w_{i,j} = q_i(c_j) - \alpha||\hat{a}_i - b_j||_2 \tag{3}$$

in which $c_j$ is the class of the $j$-th nucleus, $\hat{a}_i = a_i + \delta_i$ represents the refined position of the $i$-th anchor point, $q_i(c_j)$ is the $c_j$-th element of $q_i$, $\alpha$ is a weight term and $||\cdot||_2$ denotes $l_2$ distance. We use the Hungarian algorithm [32] to determine $\phi$ in this work. As a result, the objective of prompter is concretized as narrowing the positional and categorical difference between the selected anchor points and their matched nuclei, while ignoring the unselected anchor points as background. This objective can be achieved by minimizing the following losses.

$$\begin{aligned} \mathcal{L}_{\text{cls}} &= -\frac{1}{M}\left(\sum_{i=1}^{N} \log q_{\sigma(i)}(c_i) + \beta \sum_{a_i \in \mathcal{A}'} \log q_i(\varnothing)\right) \\ \mathcal{L}_{\text{reg}} &= \frac{\gamma}{N}\sum_{i=1}^{N} ||\hat{a}_{\sigma(i)} - b_i||_2 \end{aligned} \tag{4}$$

where $\mathcal{A}' \subsetneqq \mathcal{A}$ represents the set of unselected anchor points, $\varnothing$ indicates the background class, $\beta$ and $\gamma$ are free parameters used to relieve the class imbalance and modulate the effect of regression loss, respectively.

**Auxiliary task of nuclear region segmentation** The training process of the above prompter only involves the nuclear categorical labels and centroid coordinates. However, in the context of nucleus instance segmentation, the mask for each nucleus is also available, which provides rich details about nuclear size, shape and so on. To integrate this valuable information into prompter learning, we construct a simple auxiliary task of nuclear region segmentation to enhance the model's attention to foreground areas and perception of nuclear morphological characteristics. Technically, we introduce a mask head structured as Conv-BN-ReLu-Conv to predict the nuclear probability map $\hat{S}$ from $P_2$, informed by that the high-resolution $P_2$ contains abundant fine-grained features crucial for medical image segmentation [20]. We apply the focal loss to supervise the learning of the auxiliary task.

$$\mathcal{L}_{\text{aux}} = \text{FL}\left(\hat{S}, S\right) \tag{5}$$

where the ground-truth mask $S$ is derived from the instance map $y$ via a simple thresholding operation. The final loss used to optimize the prompter is

$$\mathcal{L}_{\text{prompter}} = \mathcal{L}_{\text{reg}} + \mathcal{L}_{\text{cls}} + \mathcal{L}_{\text{aux}} \tag{6}$$

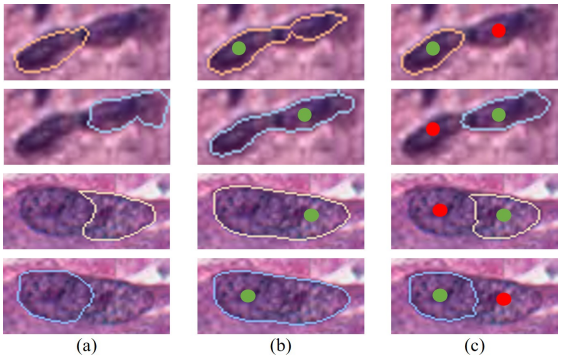


**Fig. 3:** (a) Ground-truth boundary of some overlapping nuclei. (b) Predicted boundary by prompting each nucleus with a positive prompt inside it. (c) Predicted boundary by prompting each nucleus with an additional negative prompt inside its overlapping nucleus. ● Positive prompt ● Negative prompt

**Mask-aided prompt filtering** Due to insufficient optimization, the prompter would inevitably produce false positive prompts that actually represent non-nucleus objects. To mitigate this issue, we utilize the nuclear probability map predicted by the auxiliary branch to filter out these incorrect predictions. This is achieved by retaining only those prompts with probability values exceeding 0.5 in the inference stage.

### 3.4 Use Adjacent Nuclei as Negative Prompts

Distinguishing overlapping nuclei is a long-standing challenge in the community of nucleus instance segmentation [10,12,14]. Our approach encounters this challenge as well. Given a fine-tuned SAM, considering a real-world scenario of two overlapping nuclei in Fig. 3 (a), prompting each nucleus with a single positive prompt results in an over-segmented mask due to the faint boundary, as depicted in Fig. 3 (b). An intuitive idea to resolve this problem is to include the overlapping nucleus as negative prompt to suppress excessive segmentation for the ROI nucleus, as illustrated in Fig. 3 (c).

Nevertheless, the implementation of this idea presents two practical challenges. (1) In the inference phase, it is unknown which nuclei overlap with a ROI nucleus. (2) We empirically observe that including negative prompts solely at test time cannot effectively prevent over-segmented prediction for overlapping nuclei. The inefficiency stems from that the fine-tuning process involving only positive prompts (see Eq. 1) causes a catastrophic forgetting about the effect of negative prompts.

To deal with (1), let $\hat{p}_z$ denote the generated point prompt for the $z$-th nucleus in a test image, we approximately employ the $K$ points nearest to $\hat{p}_z$ as negative prompts for segmenting this nucleus. To address (2), we incorporate negative prompts into the fine-tuning stage in a similar way. Specifically, we

randomly sample a point from each nucleus instance in $y$ and utilize the positive prompt $p_z$ along with its $K$-nearest points $\{n_{z,k}\}_{k=1}^{K}$ as negative prompts to predict the mask of the $z$-th nucleus. As a result, we re-formulate the model's forward process described by Eq. 1 as

$$\widetilde{\mathcal{O}}_z = \mathcal{M}\left(\mathcal{F}\left(x\right), \mathcal{P}\left(\{p_z\} \cup \{n_{z,k}\}_{k=1}^{K}\right), [\text{mask}], [\text{IoU}]\right) \tag{7}$$

## 4 Experiment

### 4.1 Experimental Settings

**Datasets** The experimental evaluation is conducted on three publicly available nucleus instance segmentation datasets. **Kumar** [18] consists of 30 H&E stained images (size: 1000×1000) with 21,623 annotated nuclei. The dataset is split into a training set of 16 images and a test set of 14 images. **CPM-17** [33] is comprised of 64 H&E stained images (size: 500×500 or 600×600) with 7,570 annotated nuclei. Both the training and test sets contain 32 images. **PanNuke** [8, 9] is considered one of the most challenging datasets for simultaneous nucleus instance segmentation and classification, containing 7,899 H&E stained images of 256×256 pixels and 189,744 nuclei, which are classified into five distinct classes. For this dataset, we adhere to the official three-fold cross-validation protocol [8, 9] and report the averaged results over these three splits.

**Evaluation metrics** Following preceding studies [10, 14], we employ the Aggregated Jaccard Index (AJI) and Panoptic Quality (PQ) as metrics for comparison. Given that AJI suffers from the over-penalization issue in overlapping regions [10], we designate PQ as the principal metric. For enhanced interpretability, PQ can be decomposed into two constituent parts: Detection Quality (DQ) and Segmentation Quality (SQ).

**Implementation details** are available in the appendix.

### 4.2 Comparison with SOTA Methods

We employ PromptNucSeg-B/L/H to distinguish our approach with fine-tuned SAM-B/L/H as the nucleus segmentor. Tab. 1 shows the quantitative comparison results of our approach with SOTA methods on the challenging PanNuke dataset. Without additional techniques such as stain normalization, test-time augmentation [28], oversampling or adding an auxiliary tissue classification branch [14], PromptNucSeg-H outperforms the previous best models by 1.1 bPQ and 1.4 mPQ. Moreover, we report the detection and segmentation performance of various methods for each type of nuclei in Tabs. 2 and 3. In a nutshell, our method achieves the highest F1 scores across all five classes for nucleus detection and the highest PQ scores for four out of the five categories in terms of nucleus segmentation. Tab. 4 exhibits the comparison results on the Kumar and CPM-17

**Table 1:** Performance comparison on the PanNuke dataset. Following [4, 14], both binary PQ (bPQ) and multi-class PQ (mPQ) are computed for evaluation. The best and second-best PQ scores are highlighted in **bold** and underlined.

| Tissue | Mask R-CNN [13] | | StarDist [31] | | Hover-Net [10] | | CPP-Net [4] | | PointNu-Net [37] | | CellViT-H [14] | | PromptNucSeg-H (Ours) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | bPQ | mPQ | bPQ | mPQ | bPQ | mPQ | bPQ | mPQ | bPQ | mPQ | bPQ | mPQ | bPQ | mPQ |
| Adrenal | 0.5546 | 0.3470 | 0.6972 | 0.4868 | 0.6962 | 0.4812 | 0.7066 | 0.4944 | 0.7134 | 0.5115 | 0.7086 | **0.5134** | **0.7227** | 0.5128 |
| Bile Duct | 0.5567 | 0.3536 | 0.6690 | 0.4651 | 0.6696 | 0.4714 | 0.6768 | 0.4670 | 0.6814 | 0.4868 | 0.6784 | 0.4887 | **0.6976** | **0.5012** |
| Bladder | 0.6049 | 0.5065 | 0.6986 | 0.5793 | 0.7031 | 0.5792 | 0.7053 | 0.5936 | **0.7226** | **0.6065** | 0.7068 | 0.5844 | 0.7212 | 0.6043 |
| Breast | 0.5574 | 0.3882 | 0.6666 | 0.5064 | 0.6470 | 0.4902 | 0.6747 | 0.5090 | 0.6709 | 0.5147 | 0.6748 | 0.5180 | **0.6842** | **0.5322** |
| Cervix | 0.5483 | 0.3402 | 0.6690 | 0.4628 | 0.6652 | 0.4438 | 0.6912 | 0.4792 | 0.6899 | 0.5014 | 0.6872 | 0.4984 | **0.6983** | **0.5118** |
| Colon | 0.4603 | 0.3122 | 0.5779 | 0.4205 | 0.5575 | 0.4095 | 0.5911 | 0.4315 | 0.5945 | 0.4509 | 0.5921 | 0.4485 | **0.6096** | **0.4690** |
| Esophagus | 0.5691 | 0.4311 | 0.6655 | 0.5331 | 0.6427 | 0.5085 | 0.6797 | 0.5449 | 0.6766 | 0.5504 | 0.6682 | 0.5454 | **0.6920** | **0.5711** |
| Head & Neck | 0.5457 | 0.3946 | 0.6433 | 0.4768 | 0.6331 | 0.4530 | 0.6523 | 0.4706 | 0.6546 | 0.4838 | 0.6544 | 0.4913 | **0.6695** | **0.5104** |
| Kidney | 0.5092 | 0.3553 | 0.6998 | 0.4880 | 0.6836 | 0.4424 | 0.7067 | 0.5194 | 0.6912 | 0.5066 | 0.7092 | 0.5366 | **0.7115** | **0.5786** |
| Liver | 0.6085 | 0.4103 | 0.7231 | 0.5145 | 0.7248 | 0.4974 | 0.7312 | 0.5143 | 0.7314 | 0.5174 | 0.7322 | 0.5224 | **0.7372** | **0.5333** |
| Lung | 0.5134 | 0.3182 | 0.6362 | 0.4128 | 0.6302 | 0.4004 | 0.6386 | 0.4256 | 0.6352 | 0.4048 | 0.6426 | 0.4314 | **0.6580** | **0.4398** |
| Ovarian | 0.5784 | 0.4337 | 0.6668 | 0.5205 | 0.6309 | 0.4863 | 0.6830 | 0.5313 | **0.6863** | **0.5484** | 0.6722 | 0.5390 | 0.6856 | 0.5442 |
| Pancreatic | 0.5460 | 0.3624 | 0.6601 | 0.4585 | 0.6491 | 0.4600 | 0.6789 | 0.4706 | 0.6791 | 0.4804 | 0.6658 | 0.4719 | **0.6863** | **0.4974** |
| Prostate | 0.5789 | 0.3959 | 0.6748 | 0.5067 | 0.6615 | 0.5101 | 0.6927 | 0.5305 | 0.6854 | 0.5127 | 0.6821 | 0.5321 | **0.6983** | **0.5456** |
| Skin | 0.5021 | 0.2665 | 0.6289 | 0.3610 | 0.6234 | 0.3429 | 0.6209 | 0.3574 | 0.6494 | 0.4011 | 0.6565 | **0.4339** | **0.6613** | 0.4113 |
| Stomach | 0.5976 | 0.3684 | 0.6944 | 0.4477 | 0.6886 | **0.4726** | 0.7067 | 0.4582 | 0.7010 | 0.4517 | 0.7022 | 0.4705 | **0.7115** | 0.4559 |
| Testis | 0.5420 | 0.3512 | 0.6869 | 0.4942 | 0.6890 | 0.4754 | 0.7026 | 0.4931 | 0.7058 | 0.5334 | 0.6955 | 0.5127 | **0.7151** | **0.5474** |
| Thyroid | 0.5712 | 0.3037 | 0.6962 | 0.4300 | 0.6983 | 0.4315 | 0.7155 | 0.4392 | 0.7076 | 0.4508 | 0.7151 | 0.4519 | **0.7218** | **0.4721** |
| Uterus | 0.5589 | 0.3683 | 0.6599 | 0.4480 | 0.6393 | 0.4393 | 0.6615 | 0.4794 | 0.6634 | 0.4846 | 0.6625 | 0.4737 | **0.6743** | **0.4955** |
| Average | 0.5528 | 0.3688 | 0.6692 | 0.4744 | 0.6596 | 0.4629 | 0.6798 | 0.4847 | 0.6808 | 0.4957 | 0.6793 | 0.4980 | **0.6924** | **0.5123** |
| Std | 0.0076 | 0.0047 | 0.0014 | 0.0037 | 0.0036 | 0.0076 | 0.0015 | 0.0059 | 0.0050 | 0.0082 | 0.0318 | 0.0413 | 0.0093 | 0.0147 |

**Table 2:** Precision (P), Recall (R) and F1-score (F1) for detection and classification across three folds for each nucleus type. The best F1-score is in **bold** while the second best is underlined. Following [10], if a detected nucleus is within a valid distance ($\approx 3\mu$m) from an annotated nucleus and the nuclear class matches, it is counted as a true positive (TP), otherwise a false positive(FP).

| Method | Detection | | | Classification | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Neoplastic | | | Epithelial | | | Inflammatory | | | Connective | | | Dead | | |
| | P | R | F1 | P | R | F1 | P | R | F1 | P | R | F1 | P | R | F1 | P | R | F1 |
| Mask-RCNN [13] | 0.76 | 0.68 | 0.72 | 0.55 | 0.63 | 0.59 | 0.52 | 0.52 | 0.52 | 0.46 | 0.54 | 0.50 | 0.42 | 0.43 | 0.42 | 0.17 | 0.30 | 0.22 |
| DIST [27] | 0.74 | 0.71 | 0.73 | 0.49 | 0.55 | 0.50 | 0.38 | 0.33 | 0.35 | 0.42 | 0.45 | 0.42 | 0.42 | 0.37 | 0.39 | 0.00 | 0.00 | 0.00 |
| StarDist [31] | 0.85 | 0.80 | 0.82 | 0.69 | 0.69 | 0.69 | 0.73 | 0.68 | 0.70 | 0.62 | 0.53 | 0.57 | 0.54 | 0.49 | 0.51 | 0.39 | 0.09 | 0.10 |
| Micro-Net [29] | 0.78 | 0.82 | 0.80 | 0.59 | 0.66 | 0.62 | 0.63 | 0.54 | 0.58 | 0.59 | 0.46 | 0.52 | 0.50 | 0.45 | 0.47 | 0.23 | 0.17 | 0.19 |
| Hover-Net [10] | 0.82 | 0.79 | 0.80 | 0.58 | 0.67 | 0.62 | 0.54 | 0.60 | 0.56 | 0.56 | 0.51 | 0.54 | 0.52 | 0.47 | 0.49 | 0.28 | 0.35 | 0.31 |
| CPP-Net [4] | 0.87 | 0.78 | 0.82 | 0.74 | 0.67 | 0.70 | 0.74 | 0.70 | 0.72 | 0.60 | 0.57 | 0.58 | 0.57 | 0.49 | 0.53 | 0.41 | 0.36 | 0.38 |
| CellViT-H [14] | 0.84 | 0.81 | 0.83 | 0.72 | 0.69 | **0.71** | 0.72 | 0.73 | 0.73 | 0.59 | 0.57 | 0.58 | 0.55 | 0.52 | 0.53 | 0.43 | 0.32 | 0.36 |
| PromptNucSeg | 0.82 | 0.85 | **0.84** | 0.70 | 0.72 | **0.71** | 0.73 | 0.78 | **0.76** | 0.58 | 0.61 | **0.59** | 0.55 | 0.55 | **0.55** | 0.44 | 0.49 | **0.46** |

benchmarks. In case of the Kumar dataset, our method outshines the runner-up by 0.1 points on AJI and 0.7 points on PQ. Moreover, it demonstrates a substantial improvement on the CPM-17 dataset, exceeding the second-highest AJI and PQ scores by 1.9 and 2.7 points, respectively. Fig. 4 presents the qualitative comparison results on three benchmarks.

We further analyze the model size, computational cost and inference efficiency of different methods on the PanNuke dataset in Tab. 5. The counterparts demonstrate significantly higher MACs since they generally adopt the U-Net [30]

**Table 3:** Average PQ across three folds for each nuclear category on the PanNuke dataset. The optimal results are in **bold** while the previous best arts are underlined.

| Method \ Class | Neoplastic | Epithelial | Inflammatory | Connective | Dead |
|---|---|---|---|---|---|
| Mask-RCNN [13] | 0.472 | 0.403 | 0.290 | 0.300 | 0.069 |
| DIST [27] | 0.439 | 0.290 | 0.343 | 0.275 | 0.000 |
| StarDist [31] | 0.547 | 0.532 | 0.424 | 0.380 | 0.123 |
| Micro-Net [29] | 0.504 | 0.442 | 0.333 | 0.334 | 0.051 |
| HoVer-Net [10] | 0.551 | 0.491 | 0.417 | 0.388 | 0.139 |
| CPP-Net [4] | 0.571 | 0.565 | 0.405 | 0.395 | 0.131 |
| PointNu-Net [37] | 0.578 | 0.577 | 0.433 | 0.409 | 0.154 |
| CellViT-H [14] | 0.581 | **0.583** | 0.417 | 0.423 | 0.149 |
| PromptNucSeg-H | **0.598** | 0.582 | **0.441** | **0.433** | **0.161** |

**Table 4:** Performance comparison on Kumar and CPM-17 datasets. The highest AJI and PQ scores are in **bold** while the second highest are underlined.

| Method | Kumar | | | | CPM-17 | | | |
|---|---|---|---|---|---|---|---|---|
| | AJI | DQ | SQ | PQ | AJI | DQ | SQ | PQ |
| U-Net [30] | 0.556 | 0.691 | 0.690 | 0.478 | 0.666 | 0.778 | 0.734 | 0.625 |
| DCAN [2] | 0.525 | 0.677 | 0.725 | 0.492 | 0.561 | 0.732 | 0.740 | 0.545 |
| Mask-RCNN [13] | 0.546 | 0.704 | 0.720 | 0.509 | 0.684 | 0.848 | 0.792 | 0.674 |
| DIST [27] | 0.559 | 0.601 | 0.732 | 0.443 | 0.616 | 0.663 | 0.754 | 0.504 |
| Micro-Net [29] | 0.560 | 0.692 | 0.747 | 0.519 | 0.668 | 0.836 | 0.788 | 0.661 |
| CIA-Net [41] | 0.620 | 0.754 | 0.762 | 0.577 | - | - | - | - |
| Full-Net [28] | 0.601 | 0.850 | 0.730 | 0.620 | 0.702 | 0.890 | 0.771 | 0.686 |
| Hover-Net [10] | 0.618 | 0.770 | 0.773 | 0.597 | 0.705 | 0.854 | 0.814 | 0.697 |
| Triple U-Net [40] | 0.621 | - | - | 0.601 | 0.711 | - | - | 0.685 |
| FEEDNet [7] | 0.616 | 0.843 | 0.729 | 0.613 | 0.701 | 0.894 | 0.787 | 0.705 |
| HARU-Net [3] | 0.613 | - | - | 0.572 | 0.721 | - | - | 0.701 |
| PointNu-Net [37] | 0.606 | 0.784 | 0.768 | 0.603 | 0.712 | 0.877 | 0.804 | 0.706 |
| PromptNucSeg-B | 0.614 | 0.802 | 0.773 | 0.622 | 0.731 | 0.892 | 0.813 | 0.726 |
| PromptNucSeg-L | 0.621 | 0.803 | 0.777 | 0.626 | 0.734 | 0.894 | 0.816 | 0.730 |
| PromptNucSeg-H | **0.622** | 0.803 | 0.779 | **0.627** | **0.740** | 0.897 | 0.816 | **0.733** |

architecture with progressive upsampling to regress high-resolution nuclear proxy maps. Besides, they manifest slower inference speed due to the accompanying CPU-intensive post-processing steps. In comparison, PromptNucSeg is cost-effective and efficient since it predicts nuclei prompts and their associated masks directly from hidden features of low resolution.

### 4.3 Ablation Studies

***Effect of our proposed modules*** On top of PromptNucSeg-H, we ablate the effect of our proposed modules on the CPM-17 dataset, which involve automatic nuclei prompts generation (ANPG), fine-tuning SAM (FT), auxiliary task learning of nuclear region segmentation (AUX), mask-aided prompt filtering (MAPF),

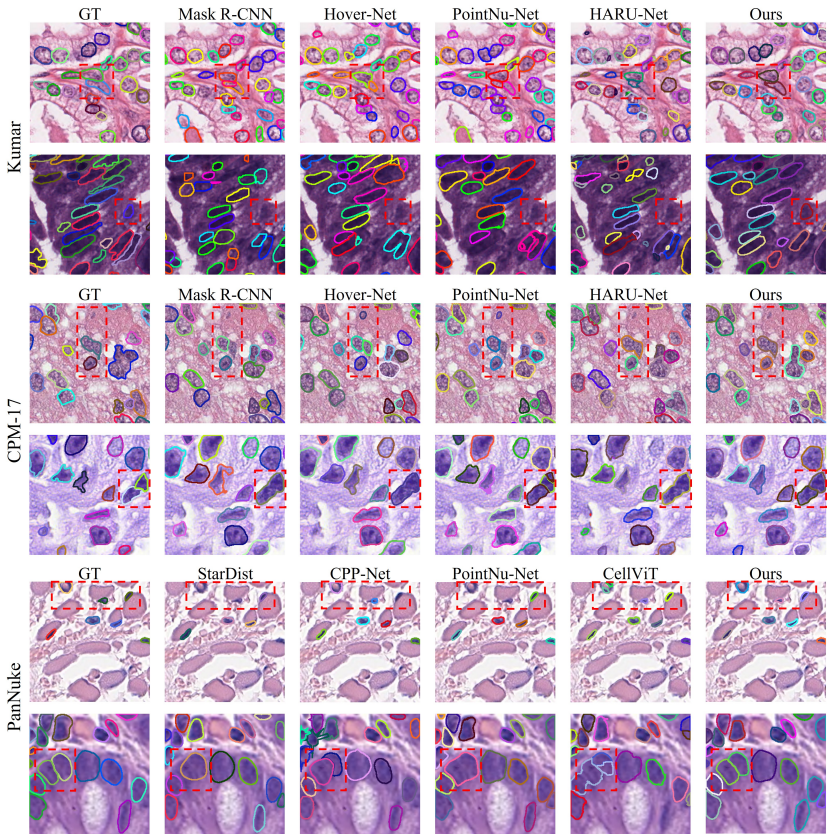


**Fig. 4:** Qualitative comparison on three benchmarks. The red dashed boxes highlight the better detection results achieved by our method.

and incorporation of negative prompts (NP). The experimental results in Tab. 6 demonstrate that all the proposed modules contribute to improving the performance of our model.

***Effect of the number of negative prompts*** We examine the impact of different quantities of negative prompts on the performance of PromptNucSeg-H using the CPM-17 dataset, as detailed in Table 7.

We initially assess the practical performance of our method by feeding predicted nuclei prompts into the segmentor. The results in Rows 1-3 discover that adding negative prompts solely in the inference stage cannot enhance the model's performance. We speculate that fine-tuning with only positive prompts results in a catastrophic forgetting about the effect of negative prompts. Comparing Rows 5 and 6, as well as Rows 8 and 9, we find that employing 1 negative prompt yields better outcomes than using 2 negative prompts. We posit that this discrepancy

**Table 5:** Comparison of model size, computational cost, efficiency and performance on the PanNuke dataset. All metrics are measured on a single NVIDIA RTX 3090 GPU.

| Method | Params (M) | MACs (G) | FPS | mPQ |
|---|---|---|---|---|
| StarDist [31] | 122.8 | 263.6 | 17 | 0.4744 |
| HoVer-Net [10] | 37.6 | 150.0 | 7 | 0.4629 |
| CPP-Net [4] | 122.8 | 264.4 | 14 | 0.4847 |
| PointNu-Net [37] | 158.1 | 335.1 | 11 | 0.4957 |
| CellViT-B [14] | 142.9 | 232.0 | 20 | 0.4923 |
| PromptNucSeg-B | 145.6 | 59.0 | 27 | 0.5095 |

**Table 6:** Effect of our proposed modules.

| ANPG | FT | AUX | MAPF | NP | AJI | PQ |
|---|---|---|---|---|---|---|
| | | | | | 0.091 | 0.091 |
| ✓ | | | | | 0.311 | 0.218 |
| | ✓ | | | | 0.496 | 0.237 |
| ✓ | ✓ | | | | 0.728 | 0.723 |
| ✓ | ✓ | | | ✓ | 0.733 | 0.727 |
| ✓ | ✓ | ✓ | | | 0.734 | 0.727 |
| ✓ | ✓ | ✓ | | ✓ | 0.737 | 0.729 |
| ✓ | ✓ | ✓ | ✓ | | 0.737 | 0.731 |
| ✓ | ✓ | ✓ | ✓ | ✓ | 0.740 | 0.733 |

**Table 7:** Effect of the number of negative prompts.

| Row | # NP | | Source of Prompts | | | |
|---|---|---|---|---|---|---|
| | Train | Test | Pred | | GT | |
| | | | AJI | PQ | AJI | PQ |
| 1 | 0 | 0 | 0.737 | 0.731 | 0.779 | 0.772 |
| 2 | 0 | 1 | 0.737 | 0.730 | 0.794 | 0.782 |
| 3 | 0 | 2 | 0.735 | 0.729 | 0.793 | 0.778 |
| 4 | 1 | 0 | 0.737 | 0.732 | 0.780 | 0.772 |
| 5 | 1 | 1 | 0.740 | 0.733 | 0.804 | 0.790 |
| 6 | 1 | 2 | 0.736 | 0.729 | 0.808 | 0.791 |
| 7 | 2 | 0 | 0.739 | 0.731 | 0.778 | 0.769 |
| 8 | 2 | 1 | 0.740 | 0.732 | 0.804 | 0.790 |
| 9 | 2 | 2 | 0.738 | 0.729 | 0.811 | 0.796 |

arises from the inherent noise in predicted prompts, the introduction of which is particularly notable when using two negative prompts.

To verify our suspicions, we further test the "oracle" performance of our method by using ground-truth nuclear centroids as prompts for the segmentor. Comparing Rows 2 and 5, we observe that the integration of negative prompts into the fine-tuning process enhances both practical and "oracle" performance. This finding confirms the existence of the catastrophic forgetting problem explained earlier. Examining Rows 4-6 and 7-9, we find that when the prompts are noise-free, the "oracle" performance continually improves with the number of negative prompts, which substantiates our second suspicion.

The substantial gaps between practical and "oracle" performance underscore the impact of prompt quality on the overall system performance. Given that training the prompter necessitates only nuclei point annotations, it is promising to improve the nucleus instance segmentation outcomes in a cost-effective scheme by bolstering the prompter's accuracy with more budget-friendly point labels.

***Which module should undertake the nucleus classification task? Answer is the prompter.*** In prior experiments, we employ the prompter for nucleus classification. Here we explore the performance of PromptNucSeg when training the prompter in a class-agnostic manner and transferring the classifi-

**Table 8:** Model performance with the prompter and segmentor as nucleus classifier, respectively.

| Classifier | Tissue | | Nucleus | | | | |
|---|---|---|---|---|---|---|---|
| | bPQ | mPQ | Neop. | Epit. | Infl. | Conn. | Dead |
| Prompter | 0.692 | 0.512 | 0.598 | 0.582 | 0.441 | 0.433 | 0.161 |
| Segmentor | 0.688 | 0.506 | 0.587 | 0.587 | 0.423 | 0.431 | 0.157 |

**Table 9:** Effect of training the prompter and segmentor jointly.

| Joint training | Kumar | | CPM-17 | | PanNuke | |
|---|---|---|---|---|---|---|
| | AJI | PQ | AJI | PQ | bPQ | mPQ |
| | 0.622 | 0.627 | 0.740 | 0.733 | 0.692 | 0.512 |
| ✓ | 0.604 | 0.605 | 0.722 | 0.710 | 0.477 | 0.352 |

cation function to the segmentor. To adapt the class-agnostic SAM for nucleus classification, we append a [cls] token to the mask decoder and update it in the same way as the [mask] and [IoU] tokens. Subsequently, the updated [cls] token is fed into a MLP head to predict the categorical logits. We incorporate a multi-class focal loss of weight 1 into Eq. 2 to supervise the classification learning. Tab. 8 displays the performance of PromptNucSeg-H on the PanNuke dataset when the prompter and segmentor are responsible for nucleus classification, respectively. The results suggest a slight performance advantage of using the prompter for nucleus classification over the segmentor.

***Does joint training lead to better performance? Answer is no.*** We explore the performance of PromptNucSeg-H when sharing the segmentor's image encoder with the prompter and training both models jointly. This adaption effectively reduces the number of model parameters by 53M but leads to a notable performance drop, as shown in Tab. 9. The performance degradation could be attributed to the limited model capacity and heightened optimization difficulty caused by hyper-parameter coupling.

## 5 Conclusion

In this paper, we have presented PromptNucSeg, a SAM-inspired method for automatic nucleus instance segmentation in histology images. Architecturally, PromptNucSeg consists of two parts: a prompter generating a distinct point prompt for each nucleus, and a segmentor predicting nuclear masks driven by these prompts. Extensive experiments across three benchmarks document the superiority of PromptNucSeg.

## Acknowledgements

This study was partially supported by the National Natural Science Foundation of China (Grant No. 92270108), Zhejiang Provincial Natural Science Foundation of China (Grant No. XHD23F0201), and the Research Center for Industries of the Future (RCIF) at Westlake University.

## References

1. Alberts, B., Bray, D., Hopkin, K., Johnson, A.D., Lewis, J., Raff, M., Roberts, K., Walter, P.: Essential cell biology. Garland Science (2015)
2. Chen, H., Qi, X., Yu, L., Heng, P.A.: Dcan: deep contour-aware networks for accurate gland segmentation. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. pp. 2487–2496 (2016)
3. Chen, J., Huang, Q., Chen, Y., Qian, L., Yu, C.: Enhancing nucleus segmentation with haru-net: A hybrid attention based residual u-blocks network. arXiv preprint arXiv:2308.03382 (2023)
4. Chen, S., Ding, C., Liu, M., Cheng, J., Tao, D.: Cpp-net: Context-aware polygon proposal network for nucleus segmentation. IEEE Transactions on Image Processing **32**, 980–994 (2023)
5. Cheng, J., Ye, J., Deng, Z., Chen, J., Li, T., Wang, H., Su, Y., Huang, Z., Chen, J., Jiang, L., et al.: Sam-med2d. arXiv preprint arXiv:2308.16184 (2023)
6. Deng, R., Cui, C., Liu, Q., Yao, T., Remedios, L.W., Bao, S., Landman, B.A., Wheless, L.E., Coburn, L.A., Wilson, K.T., et al.: Segment anything model (sam) for digital pathology: Assess zero-shot segmentation on whole slide imaging. arXiv preprint arXiv:2304.04155 (2023)
7. Deshmukh, G., Susladkar, O., Makwana, D., Mittal, S., et al.: Feednet: A feature enhanced encoder-decoder lstm network for nuclei instance segmentation for histopathological diagnosis. Physics in Medicine & Biology **67**(19), 195011 (2022)
8. Gamper, J., Alemi Koohbanani, N., Benet, K., Khuram, A., Rajpoot, N.: Pannuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. In: Digital Pathology: 15th European Congress, ECDP 2019, Warwick, UK, April 10–13, 2019, Proceedings 15. pp. 11–19. Springer (2019)
9. Gamper, J., Koohbanani, N.A., Benes, K., Graham, S., Jahanifar, M., Khurram, S.A., Azam, A., Hewitt, K., Rajpoot, N.: Pannuke dataset extension, insights and baselines. arXiv preprint arXiv:2003.10778 (2020)
10. Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T., Rajpoot, N.: Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Medical image analysis **58**, 101563 (2019)
11. He, H., Huang, Z., Ding, Y., Song, G., Wang, L., Ren, Q., Wei, P., Gao, Z., Chen, J.: Cdnet: Centripetal direction network for nuclear instance segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4026–4035 (2021)
12. He, H., Wang, J., Wei, P., Xu, F., Ji, X., Liu, C., Chen, J.: Toposeg: Topology-aware nuclear instance segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 21307–21316 (2023)
13. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision. pp. 2961–2969 (2017)

14. Hörst, F., Rempe, M., Heine, L., Seibold, C., Keyl, J., Baldini, G., Ugurel, S., Siveke, J., Grünwald, B., Egger, J., et al.: Cellvit: Vision transformers for precise cell segmentation and classification. arXiv preprint arXiv:2306.15350 (2023)
15. Huang, Y., Yang, X., Liu, L., Zhou, H., Chang, A., Zhou, X., Chen, R., Yu, J., Chen, J., Chen, C., et al.: Segment anything model for medical images? Medical Image Analysis p. 103061 (2023)
16. Ilyas, T., Mannan, Z.I., Khan, A., Azam, S., Kim, H., De Boer, F.: Tsfd-net: Tissue specific feature distillation network for nuclei segmentation and classification. Neural Networks **151**, 1–15 (2022)
17. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. arXiv preprint arXiv:2304.02643 (2023)
18. Kumar, N., Verma, R., Sharma, S., Bhargava, S., Vahadane, A., Sethi, A.: A dataset and a technique for generalized nuclear segmentation for computational pathology. IEEE transactions on medical imaging **36**(7), 1550–1560 (2017)
19. Lei, W., Wei, X., Zhang, X., Li, K., Zhang, S.: Medlsam: Localize and segment anything model for 3d medical images. arXiv preprint arXiv:2306.14752 (2023)
20. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2117–2125 (2017)
21. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. pp. 2980–2988 (2017)
22. Lin, X., Xiang, Y., Zhang, L., Yang, X., Yan, Z., Yu, L.: Samus: Adapting segment anything model for clinically-friendly and generalizable ultrasound image segmentation. arXiv preprint arXiv:2309.06824 (2023)
23. Lou, W., Wan, X., Li, G., Lou, X., Li, C., Gao, F., Li, H.: Structure embedded nucleus classification for histopathology images. arXiv preprint arXiv:2302.11416 (2023)
24. Ma, J., Wang, B.: Segment anything in medical images. arXiv preprint arXiv:2304.12306 (2023)
25. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). pp. 565–571. Ieee (2016)
26. Na, S., Guo, Y., Jiang, F., Ma, H., Huang, J.: Segment any cell: A sam-based auto-prompting fine-tuning framework for nuclei segmentation. arXiv preprint arXiv:2401.13220 (2024)
27. Naylor, P., Laé, M., Reyal, F., Walter, T.: Segmentation of nuclei in histopathology images by deep regression of the distance map. IEEE transactions on medical imaging **38**(2), 448–459 (2018)
28. Qu, H., Yan, Z., Riedlinger, G.M., De, S., Metaxas, D.N.: Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22. pp. 378–386. Springer (2019)
29. Raza, S.E.A., Cheung, L., Shaban, M., Graham, S., Epstein, D., Pelengaris, S., Khan, M., Rajpoot, N.M.: Micro-net: A unified model for segmentation of various objects in microscopy images. Medical image analysis **52**, 160–173 (2019)
30. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted

Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. pp. 234–241. Springer (2015)
31. Schmidt, U., Weigert, M., Broaddus, C., Myers, G.: Cell detection with star-convex polygons. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II 11. pp. 265–273. Springer (2018)
32. Song, Q., Wang, C., Jiang, Z., Wang, Y., Tai, Y., Wang, C., Li, J., Huang, F., Wu, Y.: Rethinking counting and localization in crowds: A purely point-based framework. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 3365–3374 (2021)
33. Vu, Q.D., Graham, S., Kurc, T., To, M.N.N., Shaban, M., Qaiser, T., Koohbanani, N.A., Khurram, S.A., Kalpathy-Cramer, J., Zhao, T., et al.: Methods for segmentation and classification of digital microscopy tissue images. Frontiers in bioengineering and biotechnology p. 53 (2019)
34. Wang, H., Guo, S., Ye, J., Deng, Z., Cheng, J., Li, T., Chen, J., Su, Y., Huang, Z., Shen, Y., et al.: Sam-med3d. arXiv preprint arXiv:2310.15161 (2023)
35. Wu, J., Fu, R., Fang, H., Liu, Y., Wang, Z., Xu, Y., Jin, Y., Arbel, T.: Medical sam adapter: Adapting segment anything model for medical image segmentation. arXiv preprint arXiv:2304.12620 (2023)
36. Xu, Q., Kuang, W., Zhang, Z., Bao, X., Chen, H., Duan, W.: Sppnet: A single-point prompt network for nuclei image segmentation. In: International Workshop on Machine Learning in Medical Imaging. pp. 227–236. Springer (2023)
37. Yao, K., Huang, K., Sun, J., Hussain, A.: Pointnu-net: Keypoint-assisted convolutional neural network for simultaneous multi-tissue histology nuclei segmentation and classification. IEEE Transactions on Emerging Topics in Computational Intelligence (2023)
38. Zhang, C., Han, D., Qiao, Y., Kim, J.U., Bae, S.H., Lee, S., Hong, C.S.: Faster segment anything: Towards lightweight sam for mobile applications. arXiv preprint arXiv:2306.14289 (2023)
39. Zhang, K., Liu, D.: Customized segment anything model for medical image segmentation. arXiv preprint arXiv:2304.13785 (2023)
40. Zhao, B., Chen, X., Li, Z., Yu, Z., Yao, S., Yan, L., Wang, Y., Liu, Z., Liang, C., Han, C.: Triple u-net: Hematoxylin-aware nuclei segmentation with progressive dense feature aggregation. Medical Image Analysis **65**, 101786 (2020)
41. Zhou, Y., Onder, O.F., Dou, Q., Tsougenis, E., Chen, H., Heng, P.A.: Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation. In: Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26. pp. 682–693. Springer (2019)