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

A Implementation Details

In fine-tuning SAM [8], we set the number Z of nuclei sampled from each training image as 25. The count K of negative prompts is set to 1, and the weight term ω in the loss function is fixed at 20. We empirically set the training batch size to 16, 32 and 64 for the Kumar, CPM-17 and PanNuke datasets, respectively.

In prompter learning, we adopt FPN [10] with ConvNeXt-S [11] backbone as the image encoder. For the PanNuke dataset, we set the level L of feature pyramid as 3 and the interval λ of pre-defined anchor points as 8 pixels. For the other two datasets, we adjust L to 4 and λ to 16 pixels. Both the regression and classification heads follow a structure of FC-ReLu-Dropout-FC. In calculating the edge values of the weighted bipartite graph, the distance term coefficient α is set to 0.1. The factor β that regulates the weight of background class in the classification loss is set to 0.2, 0.25 and 0.4 for the Kumar, CPM-17 and PanNuke benchmarks, respectively. The weight γ of the regression loss is set to 5e-3. The training batch size is empirically set as 8, 8 and 16 for the three benchmarks, respectively.

The Adam optimizer [7] with an initial learning rate of 1e-4 is employed to optimize both models. For the Kumar and CPM-17 datasets, we evaluate the performance of our method with the prompter and segmentor trained for 600 and 400 epochs, respectively. Regarding the PanNuke dataset, both the prompter and segmentor undergo training for 200 epochs, and the best-performing checkpoint on the validation set is selected for evaluation. Throughout model training, we apply the same data augmentations as [6] on the fly, which are detailed in Tab. S1.

B Generalizability Study

Building on PromptNucSeg-B, we assess the generalization capability of our method by testing the model trained with the CPM-17 train set on the Kumar

Table S1: The data augmentations are implemented with Albumentations [2].

Method	Prob.	Hyper-paramter
RandomGridShuffle	0.5	grid: (4, 4)
RandomCrop	1	height: 256
		width: 256
RandomRotate90	0.5	
HorizontalFlip	0.5	
VerticalFlip	0.5	
Downscale	0.15	scale_max: 0.5
		scale_min: 0.5
Blur	0.2	blur_limit: 10
GaussNoise	0.25	var_limit:50
ColorJitter	0.2	brightness: 0.25
		contrast: 0.25
		saturation: 0.1
		hue: 0.05
Superpixels	0.1	p_replace: 0.1
		n_segments: 200
		max_size: 128
ZoomBlur	0.1	max_factor: 1.05
RandomSizedCrop	0.1	min_max_height: (128, 256)
		height: 256
		width: 256
ElasticTransform	0.2	sigma: 25
		alpha: 0.5
		alpha-affine: 15
Normalization	1.0	mean: (0.485, 0.456, 0.406)
		std: (0.229, 0.224, 0.225)

test set, and vice versa. The experimental results are reported in Tab. S2. Clearly, our approach exhibits strong generalization to out-of-domain data, closely approaching its in-domain test results.

Following [4], we further compare the generalization ability of our method against competitors using Kumar for training and validation while combined CPM (CPM-15 and CPM-17 [16]) and TNBC [13] datasets as independent test sets. It should be noted that the Kumar samples originate from breast, liver, kidney, prostate, bladder, colon and stomach organs [9]. In contrast, the CPM samples come from brain, head and neck, and lung regions [16]. Additionally, while the TNBC dataset contains samples from an already seen organ (breast), its data is extracted from an independent source with different specimen preservation and staining practice [4]. Consequently, the experiments offer a comprehensive evaluation of our method's ability to generalize across diverse anatomical sites and data centers. The results in Table S3 demonstrate that our model has superior generalization capability compared to its counterparts.

Table S2: Generalization performance of our method. Results in parentheses indicate domain-aligned evaluation outcomes (*i.e.*, Kumar-train → Kumar-test and CPM-17-train → CPM-17-test).

	AJI	PQ
CPM-17-train → Kumar-test	0.592 (0.614)	0.598 (0.622)
Kumar-train → CPM-17-test	0.678 (0.731)	0.674 (0.726)

Table S3: Zero-shot performance comparison on the combined CPM and TNBC datasets. The highest AJI and PQ scores are in **bold** while the second highest are underlined.

Methods	Combined CPM				TNBC			
	AJI	DQ	SQ	PQ	AJI	DQ	SQ	PQ
FCN8 [12]	0.531	0.669	0.722	0.487	0.506	0.662	0.723	0.480
SegNet [1]	0.583	0.738	0.755	0.561	0.559	0.734	0.750	0.554
U-Net [15]	0.541	0.652	0.672	0.446	0.514	0.635	0.676	0.442
Mask-RCNN [5]	0.575	0.760	0.719	0.549	0.529	0.726	0.742	0.543
DCAN [3]	0.582	0.716	0.730	0.528	0.537	0.683	0.720	0.495
Micro-Net [14]	0.615	0.716	0.751	0.542	0.531	0.656	0.753	0.497
DIST [13]	0.563	0.593	0.720	0.432	0.523	0.549	0.714	0.404
HoVer-Net [4]	0.626	0.774	0.778	0.606	0.590	0.743	0.759	0.578
PointNu-Net [17]	0.586	0.771	0.778	0.603	0.564	0.765	0.775	0.594
PromptNucSeg-B	**0.655**	0.834	0.786	**0.656**	**0.608**	0.784	0.802	**0.630**

References

1. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence **39**(12), 2481–2495 (2017)
2. Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A.: Albumentations: fast and flexible image augmentations. Information **11**(2), 125 (2020)
3. 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)
4. 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)
5. 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)
6. 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)
7. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

8. 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)
9. 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)
10. 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)
11. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A convnet for the 2020s. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11976–11986 (2022)
12. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3431–3440 (2015)
13. 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)
14. 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)
15. 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)
16. 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)
17. 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)