

Co-synthesis of Histopathology Nuclei Image-Label Pairs using a Context-Conditioned Joint Diffusion Model

Seonghui Min* Hyun-Jic Oh* Won-Ki Jeong

Department of Computer Science and Engineering, Korea University, Seoul, Korea

FRI-AM-288

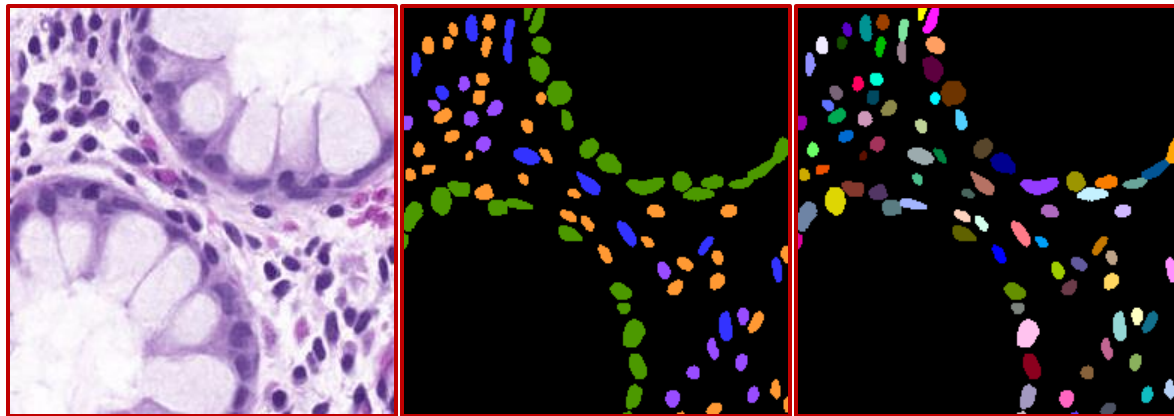
Introduction

Motivation

The limited training data becomes a bottleneck in learning-based nuclei analysis methods

→ **Solution:** Generative data augmentation to improve the model performance

Data pair for nuclei data analysis



Image

Semantic label

Instance label

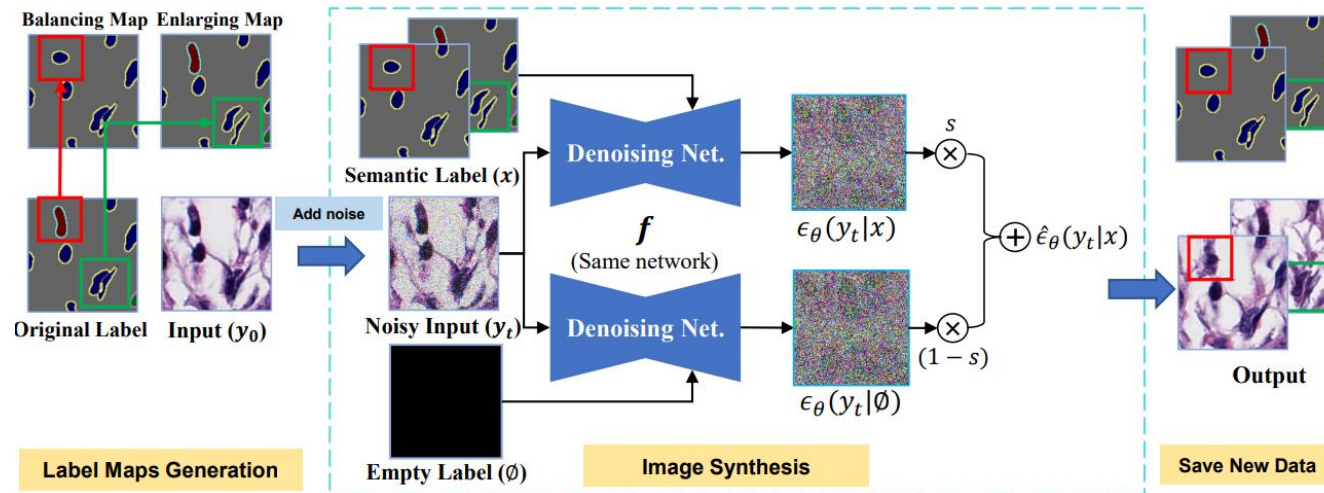
Requirements for generative data

- Realistic pathological texture
- Image-label alignment
- Accurate instance separation

Introduction

Existing work

[1] Generate images using randomly perturbed labels (e.g., copy and paste) (MICCAI23)

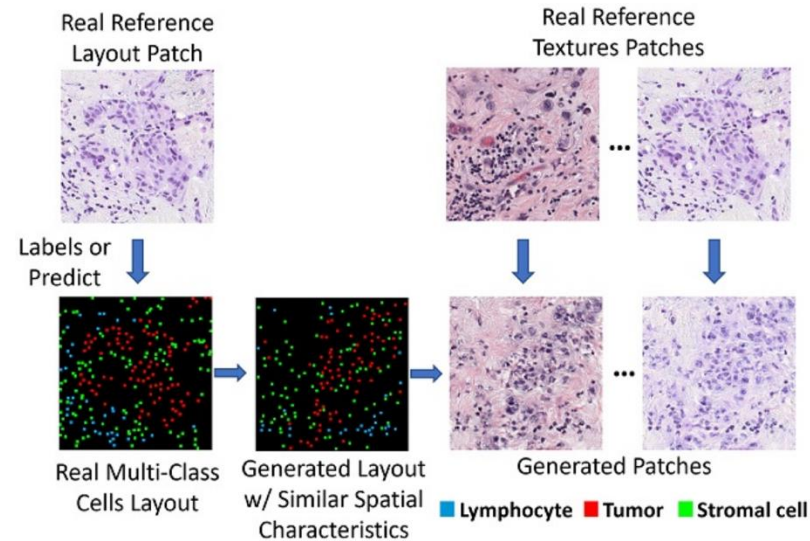


→ **Limitation:** Need to consider the pathological context

Introduction

Existing work

[2] Do not generate pixel-level semantic labels. (CVPR23)

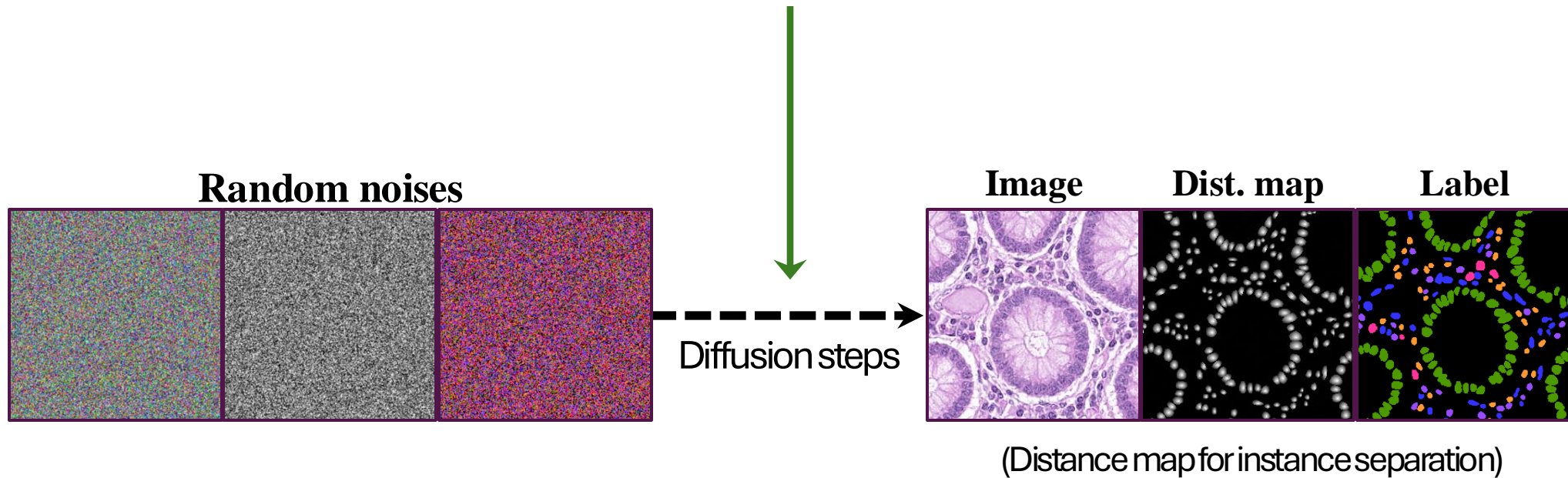


→ **Limitation:** Limited availability (not applicable for segmentation)

Introduction

Our approach: Co-synthesis image-label pair using joint diffusion model

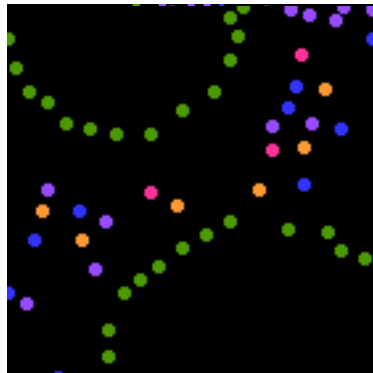
Context conditions to generate realistic data with user controllability



Introduction

Context conditions to generate realistic data with user controllability

Nucleus Centroid Layout



→ Spatially controllable

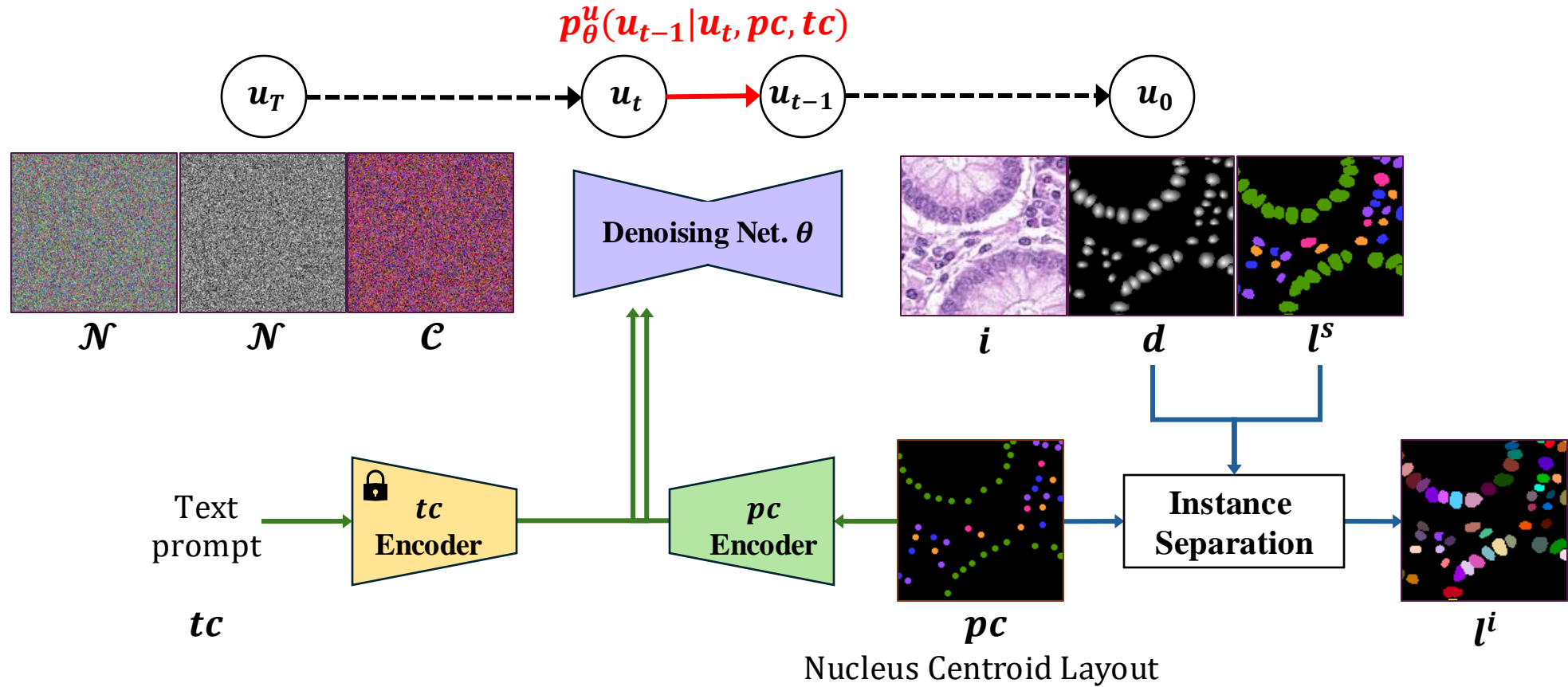
Text Prompt

*"high-quality histopathology **colon** tissue image including nuclei types of **eosinophil**, **plasma**, **lymphocyte**, **epithelial**, and **connective tissue**."*

→ Organ and nuclei types

Method

Data generation framework



→ Augment training data with synthetic image-label pair (i, l^s, l^i)

Method

Joint Diffusion Process

Forward process with normal distribution \mathcal{N} for continuous data:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I)$$

Forward process with categorical distribution \mathcal{C} for discrete data:

$$q(x_t|x_{t-1}) = \mathcal{C}(x_t; (1 - \beta_t) x_{t-1} + \beta_t/K)$$

Reverse process for data unit u , consisting of image i , distance map d , and semantic label l^s :

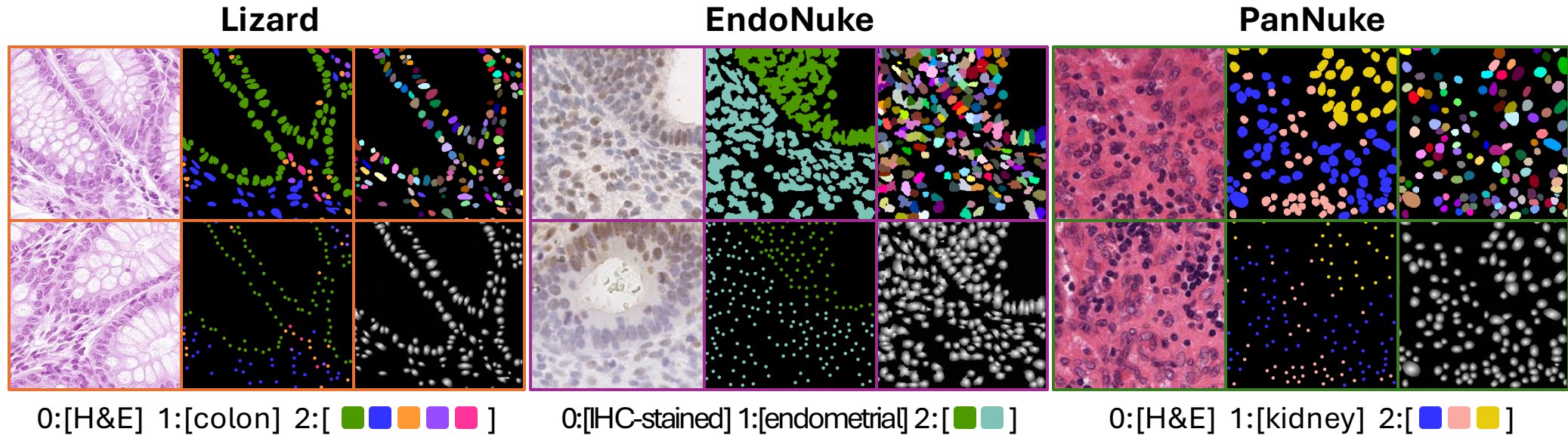
$$p_{\theta}^u(u_{t-1}|u_t) = p_{\theta}^i(i_{t-1}|u_t) \cdot p_{\theta}^d(d_{t-1}|u_t) \cdot p_{\theta}^{l^s}(l_{t-1}^s|u_t)$$

→ **Objective Function:**

$$\mathcal{L}_{total} = \lambda_i \cdot \mathcal{L}_i + \lambda_d \cdot \mathcal{L}_d + \lambda_{l^s} \cdot \mathcal{L}_{l^s}$$

Experimental results

Synthetic samples on different datasets



Syn. i	Syn. l^s	Syn. l^i
Real ref.	pc	Syn. d

tc: "... 0:[staining method] 1:[tissue type] ...2:[cell types]"

Experimental results

Quantitative comparison of generative models

Method	Lizard			PanNuke			EndoNuke		
	FID ↓	IS ↑	FSD ↓	FID ↓	IS ↑	FSD ↓	FID ↓	IS ↑	FSD ↓
Yu <i>et al.</i> [3]	-	-	963.36	-	-	1292.05	-	-	931.21
SemanticPalette [4]	86.17	2.11	<u>0.55</u>	109.23	3.36	1.23	90.0	1.40	1.88
Park <i>et al.</i> [5]	52.65	2.22	65.06	<u>61.16</u>	3.48	34.43	52.99	1.88	110.0
SDM [6]	<u>45.99</u>	<u>2.35</u>	-	107.80	3.82	-	105.17	2.27	-
Ours w/o <i>pc</i>	69.10	2.02	109.18	-	-	-	-	-	-
Ours	38.78	2.40	0.13	37.35	<u>3.77</u>	<u>1.44</u>	<u>69.94</u>	<u>2.17</u>	<u>29.57</u>

[3] Yu, X., Li, G., Lou, W., Liu, S., Wan, X., Chen, Y., Li, H.: Diffusion-based data augmentation for nuclei image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 592–602. Springer (2023)

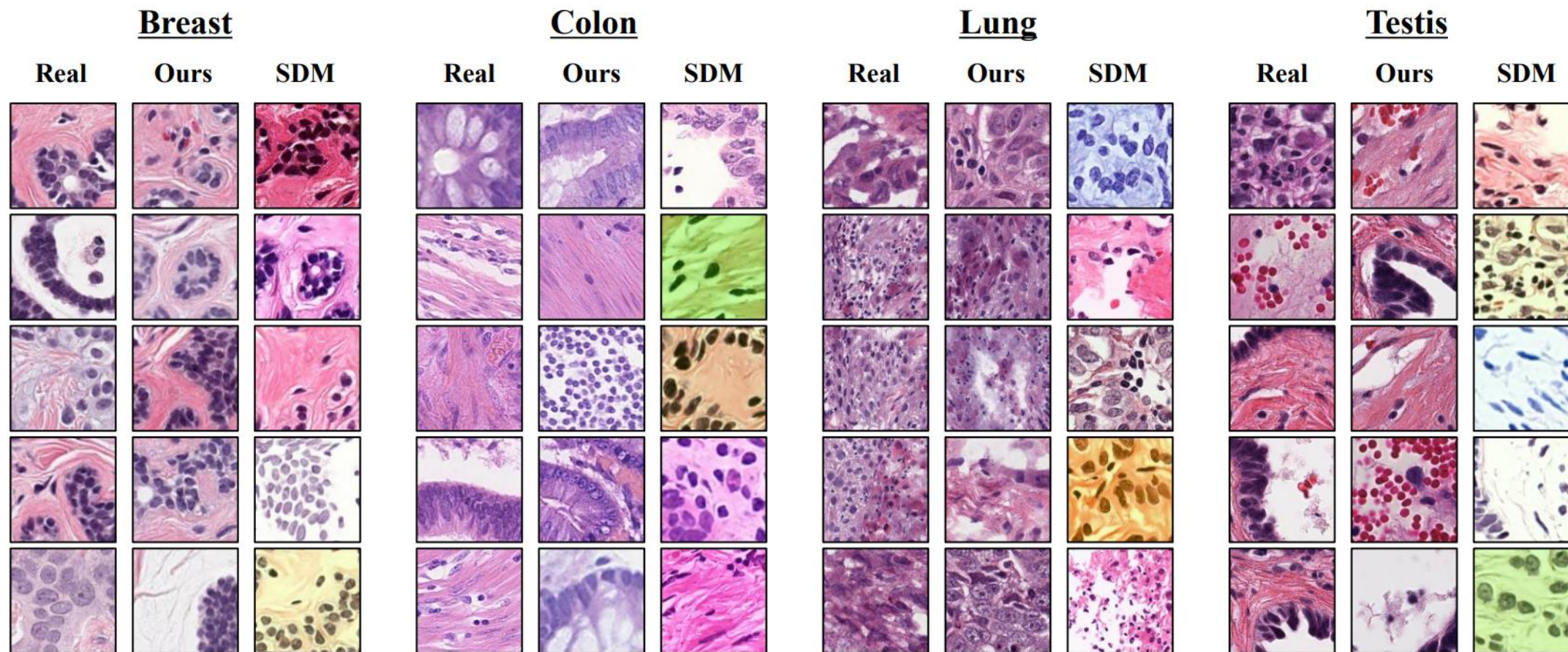
[4] Le Moing, G., Vu, T.H., Jain, H., Pérez, P., Cord, M.: Semantic palette: Guiding scene generation with class proportions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9342–9350 (2021)

[5] Park, M., Yun, J., Choi, S., Choo, J.: Learning to generate semantic layouts for higher text-image correspondence in text-to-image synthesis. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 7591–7600 (2023)

[6] Wang, W., Bao, J., Zhou, W., Chen, D., Chen, D., Yuan, L., Li, H.: Semantic image synthesis via diffusion models. arXiv preprint arXiv:2207.00050 (2022)

Experimental results

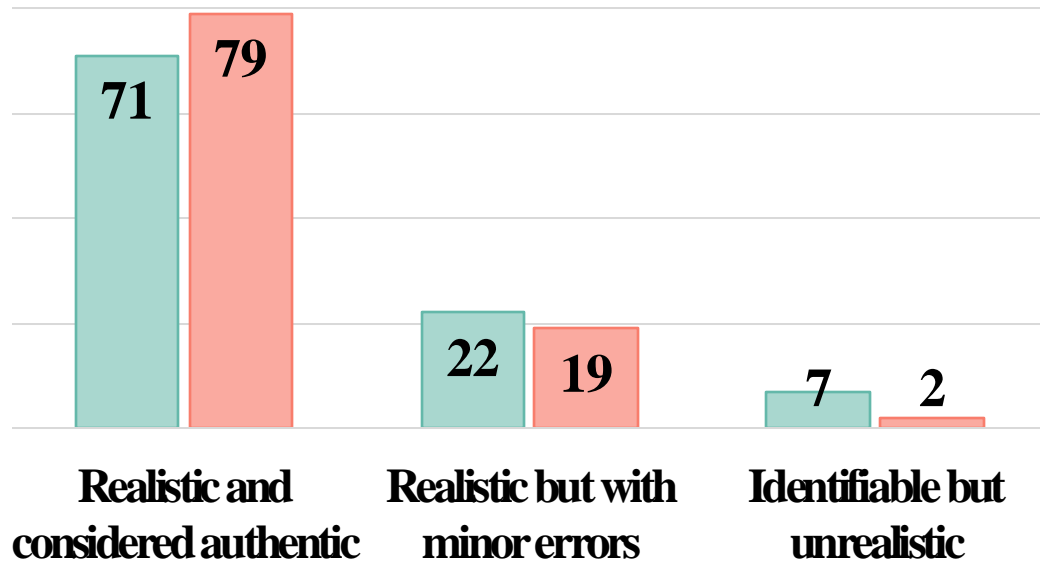
Color-quality comparison on multi-organ dataset



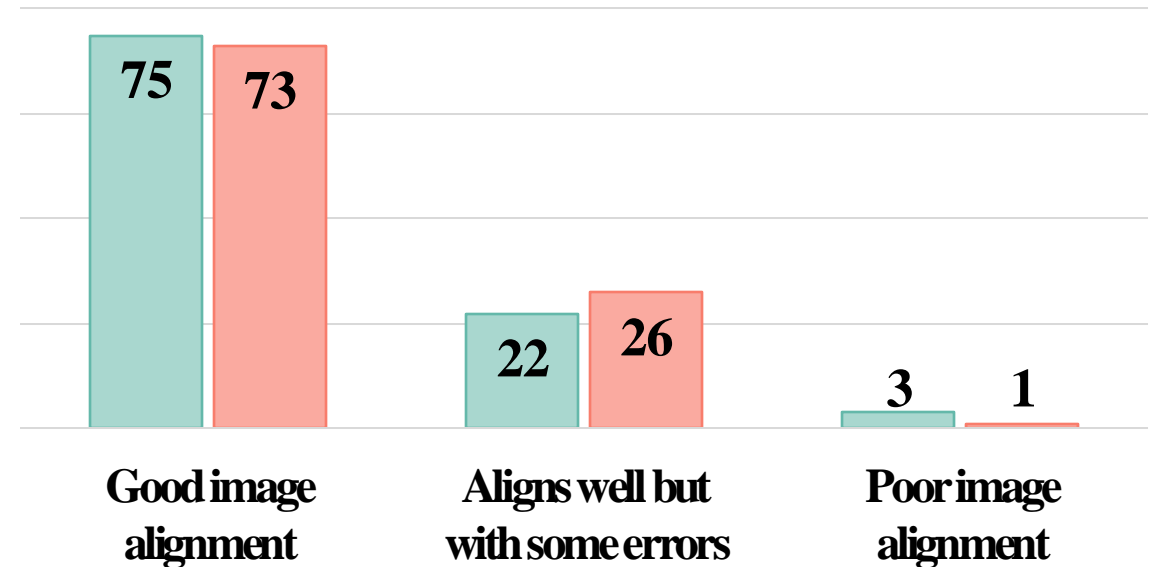
Experimental results

Qualitative assessment by pathologists

Image Assessment



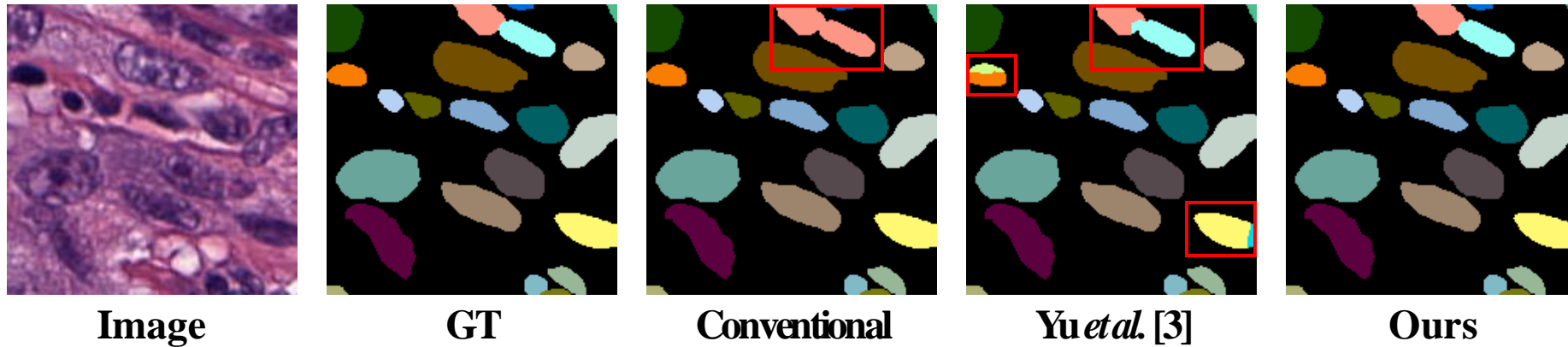
Label Assessment



Real Generated

Experimental results

Comparison of instance separation algorithms



Method	mDice		
	Lizard	PanNuke	EndoNuke
Conventional (l^S)	0.9383	0.9146	0.5524
Yu et al. [3] (d, l^S)	0.9374	0.9462	0.9268
Ours (d, l^S, pc)	0.9754	0.9980	0.9634

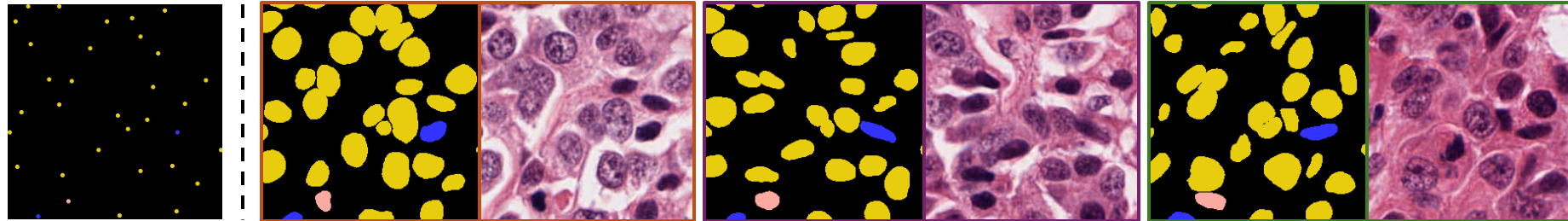
Experimental results

Downstream segmentation and classification

Dataset	Method	Segmentation		Classification	
		Dice	AJL	F_d	Acc
Lizard	Baseline	0.620	0.383	0.619	0.763
	w/ Aug	0.676	0.425	0.646	0.818
	w/ SDM [6]	0.718	0.488	0.699	<u>0.862</u>
	w/ Ours	<u>0.716</u>	<u>0.484</u>	<u>0.694</u>	0.866
PanNuke	Baseline	0.782	0.598	0.763	0.668
	w/ Aug	0.816	0.641	0.791	<u>0.708</u>
	w/ SDM [6]	<u>0.821</u>	<u>0.654</u>	<u>0.800</u>	0.702
	w/ Ours	0.824	0.662	0.806	0.736
EndoNuke	Baseline	0.878	0.594	0.815	0.891
	w/ Aug	0.889	0.602	0.820	0.905
	w/ SDM [6]	0.900	<u>0.642</u>	0.848	<u>0.909</u>
	w/ Ours	<u>0.899</u>	0.645	<u>0.844</u>	0.926

Experimental results

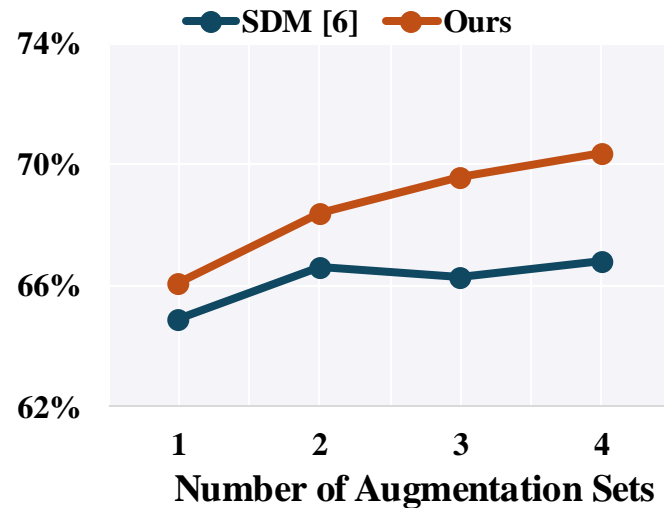
Effectiveness of diverse label generation with pc



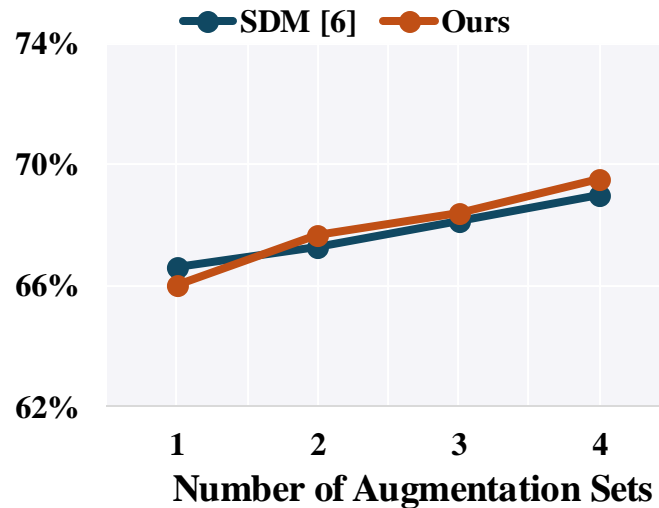
pc

Diverse Generated Image-Label Pairs

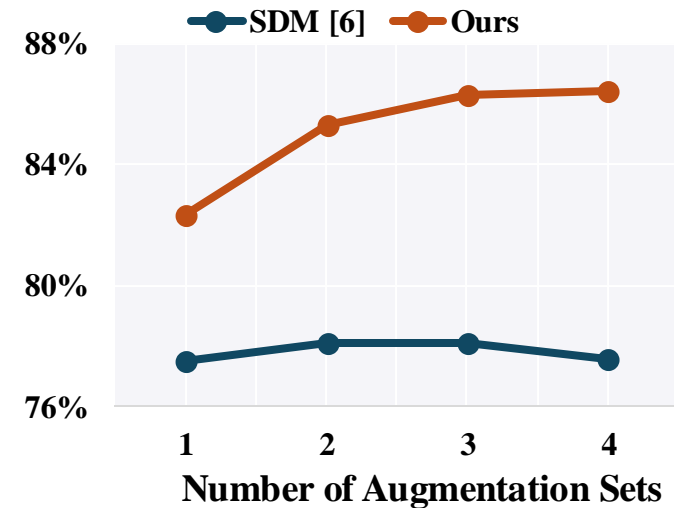
Segmentation – Dice



Classification – F_d



Classification – Accuracy



Conclusion

- A novel approach to co-synthesize image-label pairs for histopathology nuclei images
- Model the joint distribution of image, semantic label, and distance map using a single joint diffusion model
- Introduce two context conditioning methods, a point map and text prompt for precise controllable label synthesis process and realistic synthesis of histopathology images
- Use the synthesized distance mask to obtain instance label maps for downstream tasks such as nuclei instance segmentation

Future works

- Reduce the time cost for data synthesis
- Develop generative methods for creating realistic point layouts