PrPSeg: Universal Proposition Learning for Panoramic Renal Pathology Segmentation

We sincerely thank all the reviewers and ACs and the constructive feedback. In the following sections, we will address the concerns raised by the reviewers.

Evaluation with Classes Extension(R2,R3): We provide an ablation study for two data extension scenarios: (1) adding 3 new sub-types; (2) introducing 4 new objects based on reviewers comments. The proposed PrPSeg method is flexible to extend new classes by merely updating tokens and the adaptable proposition matrix, without changing the backbone network (Fig. S1). All models are trained for 30 epochs on the 15-class dataset using the same codebase and experimental settings as those described in the main manuscript. In Tab. S1, PrPSeg demonstrated superior performance compared to baseline methods across all seven new classes, maintaining the trend observed in Tab.3 of the manuscript.

Table S1. Ablation study with 7 extended classes. Dice similarity coefficient scores (%) are reported.

Method	TDH	UPL	Regions			Functional units			Cells	Mean
			Inn. Cor.	Mid. Cor.	Out. Cor.	Art.	PTC	MV	Smooth.	
Swin-Unetr			34.54	31.26	34.20	53.92	57.46	55.03	60.18	49.66
Swin-Unetr	✓		45.25	41.89	70.22	52.27	60.95	52.17	62.45	56.67
Omni-Seg			39.84	43.98	70.96	47.33	63.23	48.67	56.91	53.86
Omni-Seg	✓		51.16	46.46	70.53	57.89	62.63	61.93	64.41	60.43
PrPSeg (Ours)	✓	✓	**52.69**	**49.86**	**71.13**	**59.51**	**64.74**	**63.09**	**64.91**	**61.74**

*TDH is Token-based Dynamic Head

*UPL is Universal Proposition Learning

Novelty Clarification(R1,R2,R3): We appreciate the reviewers' comment of the innovative aspects of our method. The contributions of this paper are threefold: (1) A comprehensive universal proposition matrix is proposed to provide a simple and adaptable method to model the predominantly overlooked intricate spatial interrelations and class relationships among objects from clinical knowledge. This proposition matrix allows us to flexibly add unseen new classes via minimal changes (only modify tokens and this matrix); (2) The development of a token-based dynamic head in a single network architecture, improving partial label image segmentation. The backbone of the proposed PrPSeg network remains unchanged when new class tokens are introduced for new datasets, enabling the reuse of model weights on incomplete datasets. (3) The formulation of an anatomical loss function that quantifies the inter-object relationships across the kidney. Since the universal proposition matrix in (1) and the anatomical loss function in (3) contribute in different directions, we demonstrate them separately.

Comparison to Related Work and Contributions Beyond Medical Imaging(R1,R3): We appreciate the reviewer's suggestion for discussing several recent methods that utilize hierarchical information for semantic segmentation *Li et al., (CVPR 2022), Ke et al., (CVPR 2022)* or classification/prediction *Chen et al., (CVPR 2022)*. Our method's innovations, beyond previous work, include: (1) Emphasizing pixel-wise anatomical and spatial relation-

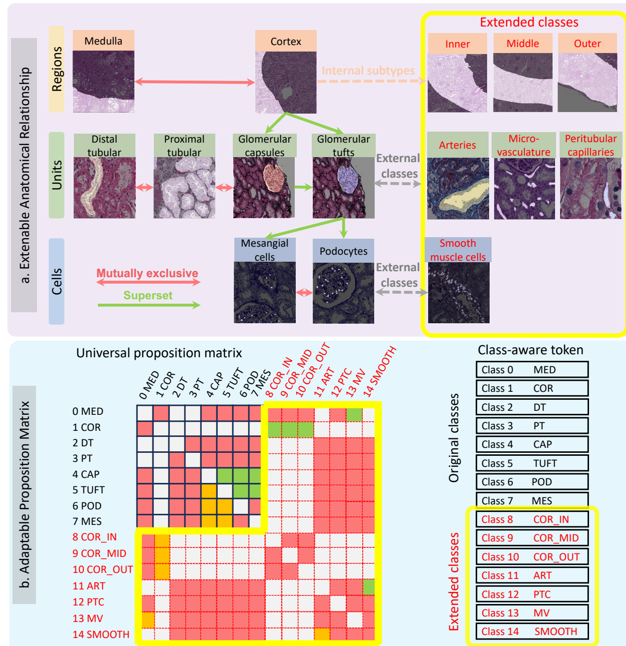

Figure S1. **The innovation of the pipeline when data is extended**

ships between objects, rather than solely taxonomy-based relationships (e.g., a glomerulus is located inside the cortex, not merely as a subset or sharing the cortex's morphology); (2) Introducing a hierarchical relationship with class tokens and scale tokens across multiple resolutions (regions at 5×, cells at 20×) to provide greater flexibility between classes and scales, as opposed to a uniform resolution in natural images; (3) Enhancing the extensibility and reusability in model design for data expansion. The proposed method aims to provide a pipeline that is scale-aware, adaptive, and anatomically aware, transitioning from the clinical domain to potential applications in incremental learning and multi-view, multi-scale learning beyond the medical field. We added suggested relative works in the manuscript revision.

Training Details for Baseline Methods(R1): All baseline models were trained using supervised learning on the entire dataset. They share the same experimental details and code-based pipeline as outlined from line 385 to line 424, except for the semi-supervised component.

Derivation of the Equation(R3): We have reformulated the equation in a more straightforward manner in Eq. (S1):

$$L_{\text{hats}}(i,j) = \begin{cases} \text{DCE}(1 - Y_i, Y_j'), & \text{if} \quad m = 1 \\ -\text{DCE}(Y_i, Y_i \cup Y_j'), & \text{if} \quad m = -1 \\ \text{DCE}(Y_i, Y_j'), & \text{if} \quad m = 2 \\ 0, & \text{if} \quad m = 0 \end{cases} \tag{S1}$$

Typos and Figure Clarification(R2, R3): We appreciate the reviewers' suggestions and will correct all typos in the final version of the manuscript.