Dynamic Graph Representation with Knowledge-aware Attention for Histopathology Whole Slide Image Analysis

Supplementary Material

1. Visualization of Patch Interaction

We visualize the dynamic interaction of patches of WiKG in Fig.1. We randomly select a patch and draw its 6 tail nodes and edges and its 6 head nodes and edges when using it as head and tail, respectively. In each epoch, the head feature is updated according to different patches, and tail information is conveyed to different patches. Note that patches containing cancers can be used as effective messages whether in macro or micro metastasis WSIs.

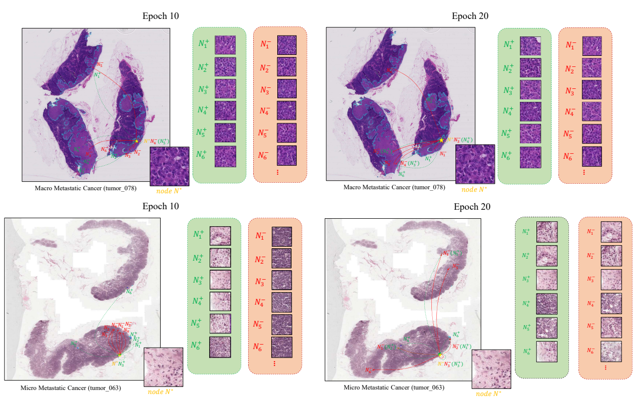

Figure 1. Changes in the outgoing and incoming edges of the selected patches at different epochs during the training phase.

2. Effect on Different Feature Extractor

To explore the impact of different feature extractors on our proposed WiKG, we comment on two other domain-specific encoders, KimiaNet[3] and SSL-DINO[2], and one ImageNet-pretrained encoder ResNet50[4]. We conduct further comparative experiments on TCGA-ESCA Cancer Typing, which is demonstrated in Table 1. The results show that WiKG performs better.

Encoder	CLAM-SB[5]		HIPT[1]		WiKG (Ours)	
	AUC	F1-score	AUC	F1-score	AUC	F1-score
ImageNet (ViT-S)	$93.56_{1.40}$	$88.28_{2.28}$	$93.79_{3.12}$	$89.27_{4.27}$	$\mathbf{95.23_{2.90}}$	$\mathbf{90.40_{3.13}}$
ImageNet (Res-50)	$92.85_{2.43}$	$86.64_{1.97}$	$86.28_{5.44}$	$78.29_{7.21}$	$\mathbf{93.44_{4.91}}$	$\mathbf{88.67_{4.44}}$
KimiaNet[3]	$95.68_{2.11}$	$89.51_{3.81}$	$93.23_{3.61}$	$89.07_{3.18}$	$\mathbf{95.78_{2.50}}$	$\mathbf{91.65_{3.12}}$
SSL-DINO[2]	$97.37_{0.97}$	$92.41_{1.03}$	$97.63_{1.53}$	$93.41_{1.24}$	$\mathbf{97.82_{1.43}}$	$\mathbf{93.75_{2.52}}$

Table 1. Comparison results of CLAM-SB, HIPT and our proposed WiKG on four different feature extractors.

3. Other Experiments of Knowledge-aware Attention

Table 2 shows the results of Knowledge-aware attention (KAA) and other GNNs on cancer staging, showing that KAA has obvious effects. Table 3 shows the difference between the maximum and minimum of k on cancer typing. It shows that Graph-based methods may not be sensitive to the neighbor node number in WSI task.

Staging	GCN	GIN	SAGE	GAT	WiKG (Ours)
ESCA	$63.77_{4.01}$	$65.30_{4.62}$	$65.55_{4.28}$	$64.37_{6.24}$	$\mathbf{69.96_{4.28}}$
LUNG	$59.19_{2.23}$	$57.82_{2.57}$	$57.78_{2.45}$	$58.47_{1.78}$	$\mathbf{60.34_{1.37}}$

Table 2. Cancer staging results of different GNN architectures

Typing	GCN	GIN	SAGE	GAT	WiKG (Ours)
ESCA	0.91	1.13	1.09	0.10	1.06
LUNG	1.09	0.84	1.53	1.41	1.49

Table 3. The maximum and minimum difference results under different numbers of neighbor nodes in the cancer typing task.

References

[1] Chen et al. Scaling vision transformers to gigapixel images via hierarchical self-supervised learning. In CVPR, 2022. 1

[2] Kang et al. Benchmarking self-supervised learning on diverse pathology datasets. In CVPR, 2023. 1

[3] Riasatian et al. Fine-tuning and training of densenet for histopathology image representation using tcga diagnostic slides. Medical Image Analysis, 2021. 1

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016. 1

[5] Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Richard Chen, Matteo Barbieri, and Faisal Mahmood. Data-efficient and weakly supervised computational pathology on whole slide images. Nature biomedical engineering, 5(6):555-570, 2021. 1