# Supplementary Material

This Supplementary Material contains the following parts:

**1. Dataset Statistics**
The summary of datasets used in this paper.

**2. Additional Results on Zero-shot WSI Tumor Subtyping**
Additional results on zero-shot WSI tumor subtyping for KEP variants pre-trained on Quilt1M and KEP variants with CNN-based visual encoders.

**3. Additional Ablation Results**
Additional ablations on the weight parameter $\alpha$ in the loss function.

**4. Discussion**
Details and discussion about the soft version of max-min positive similarity, frozen knowledge encoder, and robustness towards text prompts.

**5. Text Prompts**
The list of text prompts used in this paper.

# 1 Dataset Statistics

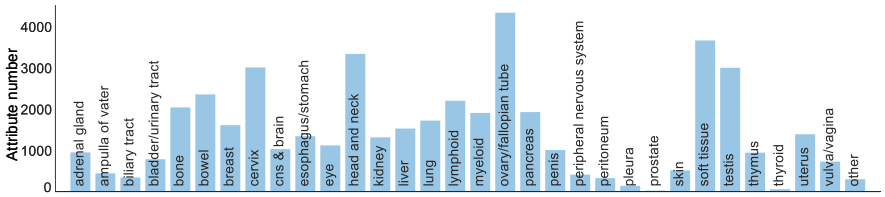


**Fig. S1:** The disease distribution in PathKT.

**Table S1:** Statistics of the dataset for knowledge encoder pretraining. Syn., Def., His. and Cyt. denote synonyms, definitions, histological and cytological features.

| Attributes | Syn. | Def. | His. | Cyt. | All |
|---|---|---|---|---|---|
| # Tissue | 32 | 32 | 30 | 30 | 32 |
| # Disease entity | 3,348 | 4,661 | 3,558 | 1,117 | 4,718 |
| # Attributes | 30,950 | 10,097 | 3,558 | 1,117 | 50,470 |

**Table S2:** Statistics of test datasets. Classification and subtyping suggest zero-shot patch-level image classification and WSI tumor subtyping, respectively.

| Task | Dataset | #images | #types | tissue |
|---|---|---|---|---|
| Retrieval | Arch-PubMed | 1923 | – | Multiple |
| | Arch-book | 1306 | – | Multiple |
| | PathPair | 9,358 | 1676 | Multiple |
| Classificaiton | BACH | 400 | 4 | Breast |
| | NCT-CRC-HE-100K | 100K | 9 | Colorectal |
| | KatherColon | 7,180 | 9 | Colon |
| | LC25000 | 25K | 5 | Lung, Colon |
| | RenalCell | 36,687 | 5 | Renal |
| | SICAP | 12,081 | 4 | Prostate |
| | SkinCancer | 129,369 | 16 | Skin |
| | WSSS4LUAD | 4,693 | 3 | Lung |
| Subtyping | TCGA-BRCA-common | 150 | 2 | Breast |
| | TCGA-BRCA-rare | 41 | 4 | Breast |
| | TCGA-NSCLC | 150 | 2 | Lung |
| | TCGA-RCC | 225 | 3 | Renal |

## 2 Additional Results on Zero-shot WSI Tumor Subtyping

In this section, we provide additional results on zero-shot WSI tumor subtyping for the KEP variants that are pretrained on Quilt1M, as shown in Fig. S2. It can be seen that KEP-16 achieves better performance than QuiltNet on BRCA and NSCLC datasets. KEP-CTP can further improve the performance, especially on the RCC dataset.

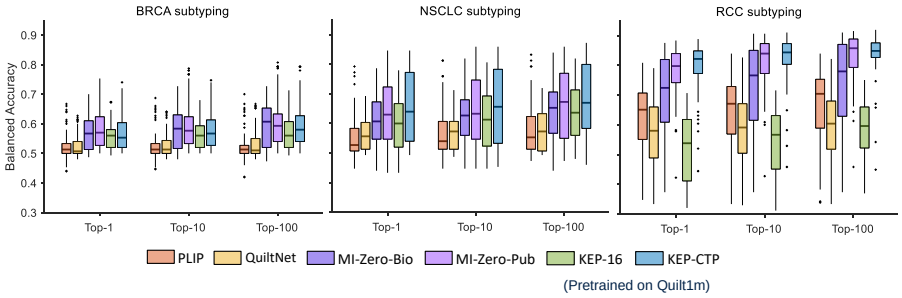


**Fig. S2:** The performance comparison of tumor subtyping on TCGA-BRCA, TCGA-NSCLC, and TCGA-RCC WSIs. The upper, center and lower line of each box denote the first, median, and third quartile of the performance distribution, respectively. The scattered points represent outliers. KEP-16 and KEP-CTP are trained on Quilt1M with the visual encoder initialized by BiomedCLIP [9] and CTransPath [7], respectively.

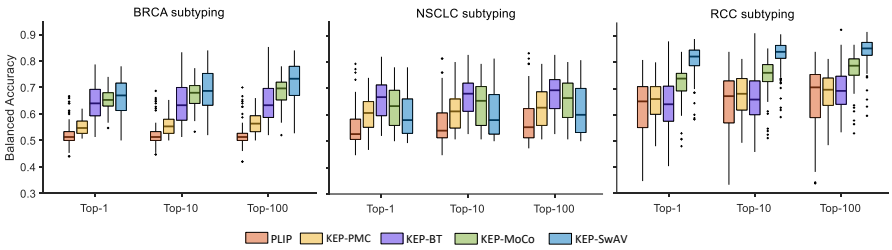


**Fig. S3:** The performance comparison of tumor subtyping on TCGA-BRCA, TCGA-NSCLC, and TCGA-RCC WSIs. The upper, center and lower line of each box denote the first, median, and third quartile of the performance distribution, respectively. The scattered points represent outliers. KEP-PMC is trained on OpenPath with the visual encoder initialized by PMC-CLIP [4], KEP-BT, KEP-MoCo, and KEP-SwAV are trained on OpenPath with the visual encoder initialized by PathSSL-BT, PathSSL-MoCo, and PathSSL-SwAV [3], respectively.

Furthermore, we change the backbone of KEP's visual encoder to convolutional neural networks and initialize it with PMC-CLIP [4] and PathSSL [3].

PMC-CLIP refers to a visual-language model pretrained on medical image-text pairs crawled from PubMed papers. PathSSL consists of PathSSL-BT, PathSSL-MoCo, and PathSSL-SwAV, which are self-supervised models pretrained on TCGA pathology images with different SSL strategies, including Barlow Twins [8], MoCo v2 [2], and SwAV [1]. Except for the visual encoder, all training details are consistent with the main text.

Fig. S3 shows the performance comparison between PLIP and different KEP variants pretrained on OpenPath. It can be seen that almost all KEP variants outperform PLIP on all datasets, suggesting that our pretraining approach can seamlessly generalize to CNN-based visual encoders with the same parameter configuration. In particular, KEP-MoCo and KEP-SwAV significantly improve the subtyping performance on breast and renal tumors.

## 3 Additional Ablation Results

Table S3 shows that KEP achieves the best performance on KatherColon dataset when the loss weight $\alpha$ is set to 0.3.

**Table S3:** Experimental results of the ablation study on the loss weight $\alpha$. Bold fonts suggest the best performance.

| $\alpha$ | 0.01 | 0.05 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 |
|---|---|---|---|---|---|---|---|
| Median | 0.531 | 0.524 | 0.531 | **0.563** | 0.556 | 0.496 | 0.481 |
| Q1 | 0.478 | 0.480 | 0.472 | **0.505** | 0.497 | 0.447 | 0.428 |
| Q3 | 0.589 | 0.587 | 0.571 | **0.610** | 0.595 | 0.540 | 0.533 |

## 4 Discussion

**Soft Version of Max-min Positive Similarity.**

$$S_i^+ = \max_p \min_q \left\langle \mathbf{z}_p^i, \mathbf{z}_q^i \right\rangle \approx \tau \log \left( \sum_{p=1}^{k} e^{\frac{\min_q \left\langle \mathbf{z}_p^i, \mathbf{z}_q^i \right\rangle}{\tau}} \right) \tag{S1}$$

$$\min_q \left\langle \mathbf{z}_p^i, \mathbf{z}_q^i \right\rangle \approx -\tau \log \left( \sum_{q=1}^{k} e^{-\frac{\left\langle \mathbf{z}_p^i, \mathbf{z}_q^i \right\rangle}{\tau}} \right) \tag{S2}$$

$$S_i^+ \approx \tau \log \left( \sum_{p=1}^{k} \frac{1}{\sum_{q=1}^{k} e^{-\frac{\left\langle \mathbf{z}_p^i, \mathbf{z}_q^i \right\rangle}{\tau}}} \right) \tag{S3}$$

**Frozen Knowledge Encoder.** Here, we discuss the advantages of the frozen knowledge encoder. Although the visual-language pretraining can start in a well-defined pathology embedding space curated by the knowledge initialization of

the text encoder, the free text from the training pairs may impair the knowledge structure built from the pathology knowledge tree, leading to overfitting on the training image-text pairs. As a consequence, the frozen knowledge encoder across the entire training period holds two advantages: (i) the text embedding from this branch acts as a frozen knowledge continuously distilling to the active text encoder, which, therefore, can substantially keep the entire alignment procedure in the well-structured knowledge embedding space; (ii) the additional contrastive loss between the active text encoder and this branch serves as a regularization term to prevent the overfitting problem.

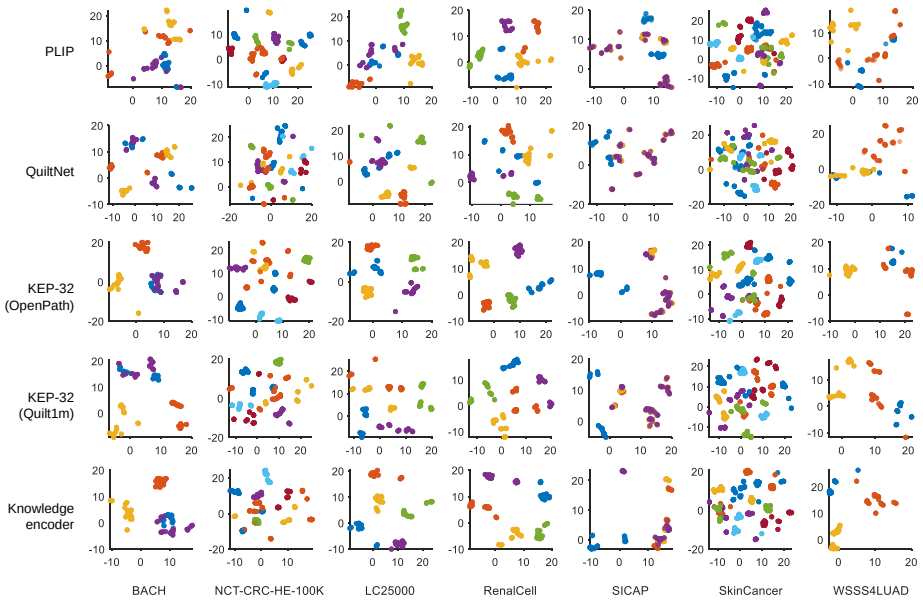


**Fig. S4:** The UMAP visualization of prompt embeddings on different datasets. Each color represents a class of the dataset and consists of 100 points corresponding to the random 100 prompts for each category. NCT-CRC-HE-100K and KatherColon share the same text prompts. KEP-32 (OpenPath) and KEP-32 (Quilt1M) denote the prompt embeddings generated from the text encoder pretrained on the dataset of OpenPath and Quilt1M, respectively. Knowledge encoder denotes the prompt embeddings from the frozen knowledge encoder.

**Robustness towards Text Prompts.** To show the robustness of our approach towards different text prompts, we use UMAP [6] to visualize the prompt embeddings for each class, as shown in Fig. S4. It can be concluded that the prompt embeddings of PLIP and QuiltNet suffer from high similarities between different classes, especially in the dataset of BACH, LC25000, SICAP, and WSSS4LUAD, which therefore causes large ambiguities for zero-shot classification. In contrast, the embeddings generated by our proposed frozen knowledge encoder (denoted

as Knowledge encoder), are well-separated in most datasets, With the knowledge guidance, the text encoder of KEP, denoted by KEP-32, also outputs well-separated prompt embeddings of different classes, suggesting that pathology knowledge can substantially improve the structure of text embedding space.

## 5 Text Prompts

In this section, we list all text prompts used in this paper. The templates are shown in Tab. S4. The synonyms for the patch-level dataset are shown in Tab. S5-S11. The class names for TCGA WSIs are exhibited in Tab. S12-S14.

**Table S4:** Prompt templates used in this paper, which are consistent with CONCH [5], CLASSNAME is replaced by the names/synonyms of classes.

| |
|---|
| CLASSNAME. |
| a photomicrograph showing CLASSNAME. |
| a photomicrograph of CLASSNAME. |
| an image of CLASSNAME. |
| an image showing CLASSNAME. |
| an example of CLASSNAME. |
| CLASSNAME is shown. |
| this is CLASSNAME. |
| there is CLASSNAME. |
| a histopathological image showing CLASSNAME. |
| a histopathological image of CLASSNAME. |
| a histopathological photograph of CLASSNAME. |
| a histopathological photograph showing CLASSNAME. |
| shows CLASSNAME. |
| presence of CLASSNAME. |
| CLASSNAME is present. |
| an H&E stained image of CLASSNAME. |
| an H&E stained image showing CLASSNAME. |
| an H&E image showing CLASSNAME. |
| an H&E image of CLASSNAME. |
| CLASSNAME, H&E stain. |
| CLASSNAME, H&E. |

**Table S5:** Class names of BACH.

| Class | Names/Synonyms |
|---|---|
| Benign | breast non-malignant benign tissue; breast benign tissue; non-malignant benign tissue of breast |
| InSitu | breast malignant in-situ carcinoma; breast in-situ carcinoma; malignant carcinoma in-situ of breast |
| Invasive | breast malignant invasive carcinoma; breast invasive carcinoma; invasive carcinoma of breast |
| Normal | normal breast tissue; breast normal tissue; breast non-cancerous tissue |

**Table S6:** Class names of NCT-CRC-HE-100K and KatherColon from CONCH [5].

| Class | Names/Synonyms |
|---|---|
| ADI | adipose; adipose tissue; adipocytes; fat; fat cells |
| BACK | background; penmarking; empty space; background artifacts |
| DEB | debris; colorectal adenocarcinoma debris and necrosis; necrosis; necrotic debris |
| LYM | lymphocytes; lymphoid aggregate; immune cells; lymphoid infiltrate; inflammatory cells |
| MUC | mucus; mucin; mucus pool; mucin pool |
| MUS | smooth muscle; smooth muscle tissue; muscle; muscularis propria; muscularis mucosa |
| NORM | normal colon mucosa; uninvolved colon mucosa; benign colon mucosa; benign epithelium |
| STR | cancer-associated stroma; tumor-associated stroma; stromal cells; stromal tissue; stroma |
| TUM | colorectal adenocarcinoma epithelium; colorectal adenocarcinoma; tumor; adenocarcinoma; malignant epithelium |

**Table S7:** Class names of LC25000.

| Class | Names/Synonyms |
|---|---|
| lung_aca | lung adenocarcinoma; adenocarcinoma of the lung; lung cancer, adenocarcinoma |
| lung_n | benign lung; benign lung tissues; non-malignant lung tissue |
| lung_scc | lung squamous cell carcinoma; squamous-cell carcinoma of the lung; squamous cell lung cancer |
| colon_aca | colon adenocarcinoma; adenocarcinoma of the colon; colon cancer, adenocarcinoma |
| colon_n | benign colon; benign colonic tissue; non-malignant colon tissue |

**Table S8:** Class names of RenalCell.

| Class | Names/Synonyms |
|---|---|
| blood | red blood cells; red blood corpuscles; red cells; erythroid cells |
| cancer | non-tumor; normal tissue; non-cancerous tissue |
| normal | renal cancer; renal tumor; renal neoplasm; renal carcinoma |
| other | torn adipose necrotic tissue; torn adipose tissue, necrosis; adipose necrotic tissue |
| stroma | muscle fibrous stroma blood vessels; blood vessels, muscle fibrous stroma; muscle fibers and blood vessels in stroma |

**Table S9:** Class names of SICAP from CONCH [5].

| Class | Names/Synonyms |
|---|---|
| NC | non-cancerous tissue; non-cancerous prostate tissue; benign tissue; benign glands; benign prostate tissue; benign prostate glands |
| G3 | gleason grade 3; gleason pattern 3; prostate cancer, gleason grade 3; prostate cancer, gleason pattern 3; prostate adenocarcinoma, well-differentiated; well-differentiated prostatic adenocarcinoma |
| G4 | gleason grade 4; gleason pattern 4; prostate cancer, gleason grade 4; prostate cancer, gleason pattern 4; prostate adenocarcinoma, moderately differentiated; moderately differentiated prostatic adenocarcinoma |
| G5 | gleason grade 5; gleason pattern 5; prostate cancer, gleason grade 5; prostate cancer, gleason pattern 5; prostate adenocarcinoma, poorly differentiated; poorly differentiated prostatic adenocarcinoma |
| Tumor | prostatic adenocarcinoma; adenocarcinoma; prostate cancer; tumor tissue; cancerous tissue |

**Table S10:** Class names of WSSS4LUAD from CONCH [5].

| Class | Names/Synonyms |
|---|---|
| normal | non-tumor; normal tissue; non-cancerous tissue |
| stroma | tumor-associated stroma; cancer-associated stroma; tumor-associated stromal tissue; cancer-associated stromal tissue |
| tumor | tumor tissue; tumor epithelial tissue; cancerous tissue |

**Table S11:** Class names of SkinCancer.

| Class | Names/Synonyms |
|---|---|
| necrosis | necrosis; necrotic tissue; necrotic cells |
| skeletal | skeletal muscle; skeletal muscle cells; skeletal muscle tissue |
| sweatglands | eccrine sweat glands; merocrine glands; skin eccrine sweat glands |
| vessel | vessels; blood vessels; vessel |
| elastosis | elastosis; elastosis of skin; skin elastosis |
| chondraltissue | chondral tissue; chondral tissue of skin; skin chondral tissue |
| hairfollicle | hair follicle; hair follicle of skin; skin hair follicle |
| epidermis | epidermis; skin epidermis; epidermal cells |
| nerves | nerves; nerve fibers; nerve axons |
| subcutis | subcutis; subcutaneous tissue; skin subcutis; hypodermis; hypoderm |
| dermis | dermis; skin dermis; corium; skin corium |
| sebaceousglands | sebaceous; sebaceous gland; skin sebaceous |
| sqcc | squamous-cell carcinoma; cutaneous squamous-cell carcinoma; squamous-cell carcinoma of the skin; squamous-cell skin cancer |
| melanoma | melanoma in-situ; malignant melanoma; cutaneous melanoma |
| bcc | basal-cell carcinoma; basal-cell cancer; basal-cell tumor |
| naevus | naevus; mole; skin nevus |

**Table S12:** Class names of TCGA-BRCA, which are consistent with CONCH [5].

| Class | Names/synonyms |
|---|---|
| IDC | invasive ductal carcinoma; breast invasive ductal carcinoma; invasive ductal carcinoma of the breast; invasive carcinoma of the breast, ductal pattern; breast IDC |
| ILC | invasive lobular carcinoma; breast invasive lobular carcinoma; invasive lobular carcinoma of the breast; invasive carcinoma of the breast, lobular pattern; breast ILC |

**Table S13:** Class names of TCGA-NSCLC, which are consistent with CONCH [5].

| Class | Names/synonyms |
|---|---|
| LUAD | adenocarcinoma; lung adenocarcinoma; adenocarcinoma of the lung; LUAD |
| LUSC | squamous cell carcinoma; lung squamous cell carcinoma; squamous cell carcinoma of the lung; LUSC |

**Table S14:** Class names of TCGA-RCC, which are consistent with CONCH [5].

| Class | Names/synonyms |
|---|---|
| CCRCC | clear cell renal cell carcinoma; renal cell carcinoma, clear cell type; renal cell carcinoma of the clear cell type; clear cell RCC |
| PRCC | papillary renal cell carcinoma; renal cell carcinoma, papillary type; renal cell carcinoma of the papillary type; papillary RCC |
| CHRCC | chromophobe renal cell carcinoma; renal cell carcinoma, chromophobe type; renal cell carcinoma of the chromophobe type; chromophobe RCC |

## References


1. Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., Joulin, A.: Unsupervised learning of visual features by contrasting cluster assignments. Advances in Neural Information Processing Systems (NeurIPS) **33**, 9912–9924 (2020)
2. Chen, X., Fan, H., Girshick, R., He, K.: Improved baselines with momentum contrastive learning. arXiv preprint arXiv:2003.04297 (2020)
3. Kang, M., Song, H., Park, S., Yoo, D., Pereira, S.: Benchmarking self-supervised learning on diverse pathology datasets. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3344–3354 (2023)
4. Lin, W., Zhao, Z., Zhang, X., Wu, C., Zhang, Y., Wang, Y., Xie, W.: Pmc-clip: Contrastive language-image pre-training using biomedical documents. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) (2023)
5. Lu, M.Y., Chen, B., Williamson, D.F., Chen, R.J., Liang, I., Ding, T., Jaume, G., Odintsov, I., Zhang, A., Le, L.P., et al.: Towards a visual-language foundation model for computational pathology. arXiv preprint arXiv:2307.12914 (2023)
6. McInnes, L., Healy, J., Melville, J.: Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018)
7. Wang, X., Yang, S., Zhang, J., Wang, M., Zhang, J., Yang, W., Huang, J., Han, X.: Transformer-based unsupervised contrastive learning for histopathological image classification. Medical Image Analysis **81**, 102559 (2022)
8. Zbontar, J., Jing, L., Misra, I., LeCun, Y., Deny, S.: Barlow twins: Self-supervised learning via redundancy reduction. In: International Conference on Machine Learning (ICML). pp. 12310–12320. PMLR (2021)
9. Zhang, S., Xu, Y., Usuyama, N., Bagga, J., Tinn, R., Preston, S., Rao, R., Wei, M., Valluri, N., Wong, C., et al.: Large-scale domain-specific pretraining for biomedical vision-language processing. arXiv preprint arXiv:2303.00915 (2023)