Leveraging Image Captions for Selective Whole Slide Image Annotation

Jingna Qiu[1], Marc Aubreville[2,3], Frauke Wilm[4], Mathias Öttl[4], Jonas Utz[1], Maja Schlereth[1], and Katharina Breininger[1,5]

[1] Department Artificial Intelligence in Biomedical Engineering,
FAU Erlangen-Nürnberg, Erlangen, Germany `jingna.qiu@fau.de`
[2] Technische Hochschule Ingolstadt, Ingolstadt, Germany
[3] Flensburg University of Applied Sciences, Flensburg, Germany
[4] Pattern Recognition Lab, Department of Computer Science,
FAU Erlangen-Nürnberg, Erlangen, Germany
[5] Center for AI and Data Science, Universität Würzburg, Würzburg, Germany
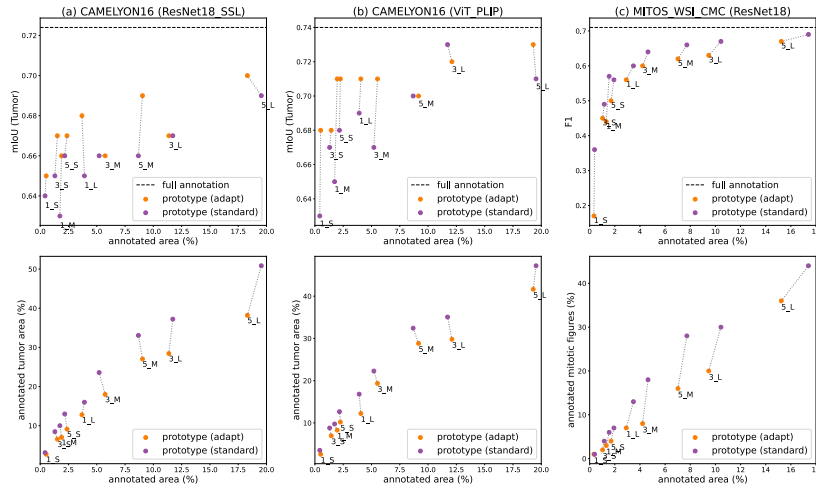
1 Supplementary Materials

Fig. S1: Comparison of the standard and adaptive region selection method. (a-b) Results on CAMELYON16 dataset: mIoU (Tumor) and annotated tumor area (%) as functions of annotated tissue area (%) for prototype sampling across various hyperparameter settings. (c) Results on MITOS_WSI_CMC dataset: F1 and the ratio of annotated mitotic figures as functions of annotated tissue area (%) for prototype sampling across various hyperparameter settings. All results show the median value from five repetitions.

Table S1: Pre- and post-processing steps for obtaining class prototype images by keyword search from the ARCH database.

	Method	Examples
Pre-processing	1. Subcaptions division by pattern recognition (the subfigures have been individually split by the data supplier, each accompanied by the original caption). Before image search, we divide subcaptions by recognizing patterns. 2. Any supcaption that is applied to all subfigures are appended to each subcaption.	We used the following patterns: "A, xxx.", "(A) xxx.", and "xxx (A).", where A and xxx represent the subfigure identifier and the subcaption.
Post-processing	Manually clean the recognized images.	We removed an image with caption containing "melanoma metastatic to the breast", it was recognized as containing keywords "breast" and "metastatic". We removed an image with caption containing 'lack mitotic activity", it was recognized as containing keywords 'mitotic".

Table S2: Our method uses class names as keywords for prototype image retrieval, under the assumption that data analysis experts lack medical expertise. This design may lead to an incomplete retrieval of prototype images. We share several pitfalls we noticed during experiments and our solutions/suggestions for promoting future investigations.

Problem	Example	Solution/Suggestion
The class name is not included in the caption.	The phrase "SNL micrometastasis possibly derived from breast metastasis" cannot be located using the search term "breast cancer".	We alleviated this type of problems by offering synonyms.
Complex medical terminology is employed to describe the condition without explicitly stating the name of the disease.	"Ductal carcinoma in situ" refers to a specific type of breast cancer.	Having a list of medical terms offered by specialists is beneficial.
Contextual information in the caption hinders accurate image recognition.	A huge nodal melanocytic nevus image was detected while searching for breast cancer images because the caption contains "Incidental large nevus found during an axillary node dissection in a patient with breast carcinoma.".	We resolved this type of error by manually cleaning irrelevant images.