CVPR
#9472

CVPR
#9472

CVPR 2023 Submission #9472. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

A Loopback Network for Explainable Microvascular Invasion Classification
—— *Supplementary Material* ——

Anonymous CVPR submission

Paper ID 9472

In the supplementary materials, we provide details of the correlation filter, visualization results, and performance on pseudo mask iteration. In addition, source codes are given in the '*source_codes.zip*' file.

A. Details of Correlation Filter

In order to distinguish cells and backgrounds, we utilize the kernelized correlation filter [2, 4] to generate pseudo masks of cells in vessel images. According to the cell appearance, we manually select some cells to generate correlation filter kernels, as shown in Fig.1. As mentioned in [4], the filter kernel $\mathbf{w}$ of each cell template is calculated as follows:

$$\hat{\mathbf{w}} = \frac{\hat{\mathbf{x}}^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda},$$ (1)

where $\mathbf{x}$ is cell image features extracted by Histogram of Oriented Gradients (HOG) [3] and Color Name (CN) [7], $\mathbf{x}^*$ denotes the complex-conjugate of $\mathbf{x}$, $\hat{x} = DFT(x)$ is a shorthand for the discrete Fourier transform, $\odot$ is defined as element-wise product, λ is a regularization parameter and we set it to 0.001, $\mathbf{y}$ is a Gaussian shaped regression target.

In order to adapt the different sizes of cells, we resize the square cell templates to slide lengths 36, 52, 68, 84, and 100. The HOG and CN features are extracted in $4\times$ downsample, and their channels are 9 and 11, respectively. We use the same approach to extract features from input vessel images and perform the correlation filter with the cell template kernels. The mean response of the top 3 high kernels is regarded as the final result. For every template kernel, the mean response of the top 3 high sizes is the response of the kernel. Each local maximum in the final result is treated as the centroid of a cell. Within every circle with a radius of r centered at the cell centroid, the pixels whose values are higher than t times the center value are regarded as the pseudo mask of the cell. The t and r are set to 0.88 and 25, respectively. Some pseudo mask examples are shown in Fig.2.
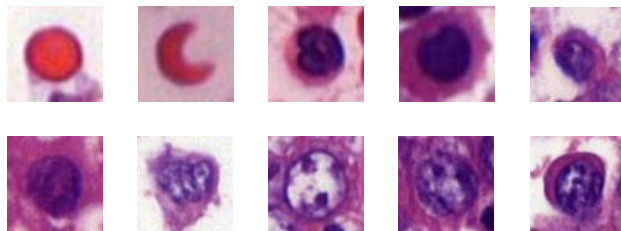

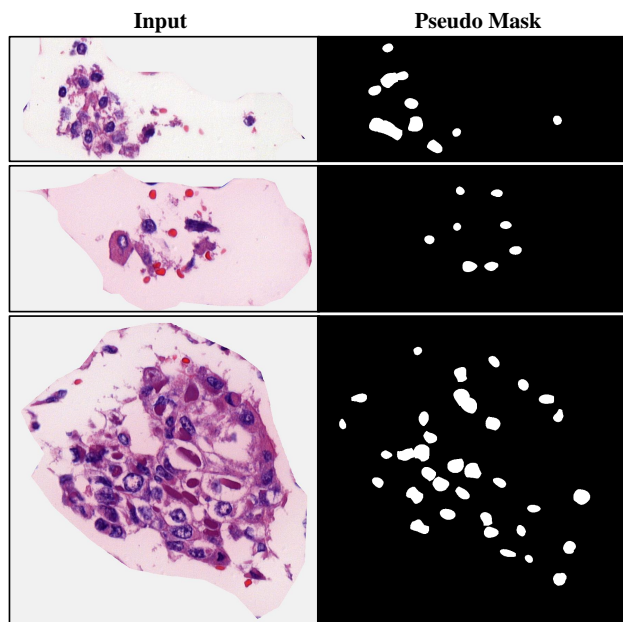Figure 1. Examples of cells for generating correlation filter kernels.

Figure 2. Examples of cell pseudo masks generated by the correlation filter.

B. Visualization Results

The visualization results of LoopNet and other feature attribution method mentioned in the main text is shown in Fig.3. For fair comparison, we only visualize cancerous cell patches, rather than the pixel-wise results, predicted by ev-
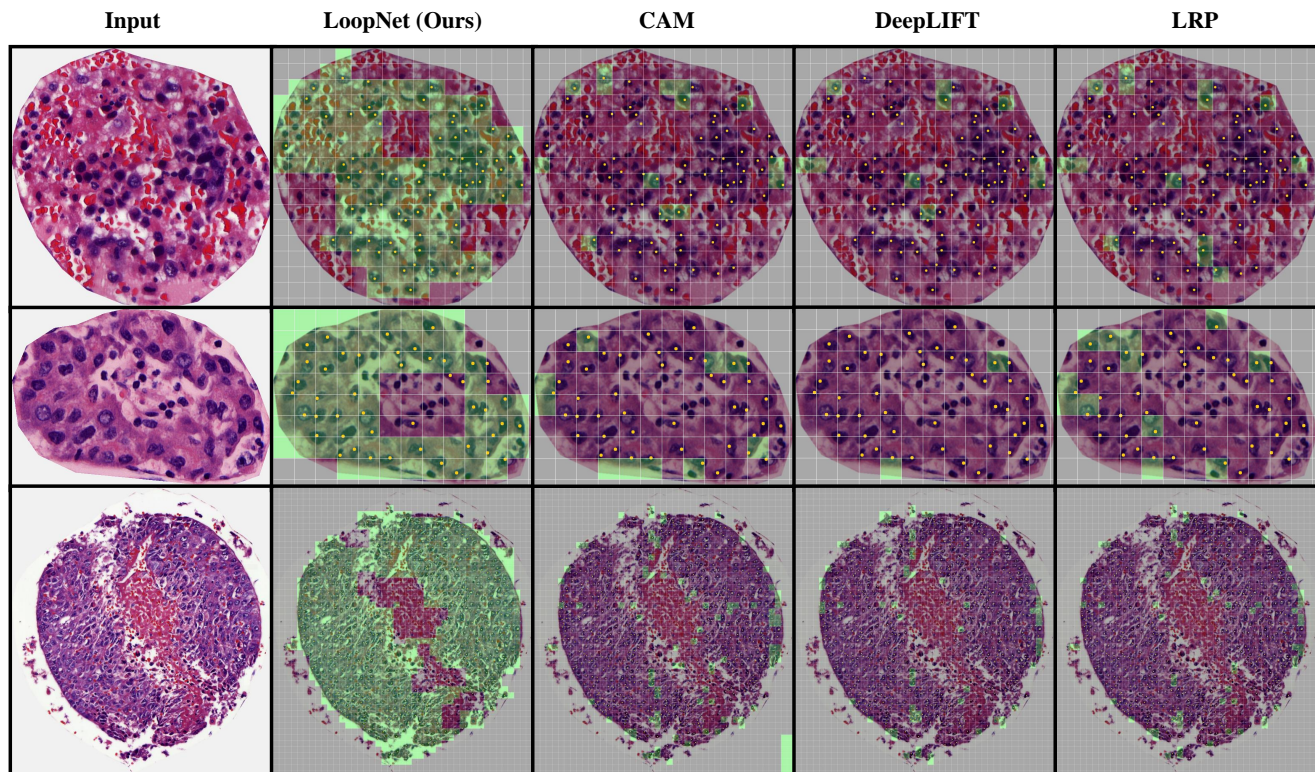
Figure 3. The visualization results of LoopNet and other feature attribution methods. We only visualize cancerous cell patches predicted by every approach using green color for a fair comparison, and the cancerous cell centroid annotations labeled by the pathologist are denoted as orange points in each result image.

ery approach using green color. The cancerous cell centroids labeled by the pathologist are noted as orange points in the result images. For CAM [5], we set the target layer to the last convolution layer of the backbone $\mathcal{F}_b$, and the positions of the class activation map for the MVI category where the value is greater than 0.5 is marked in green. For DeepLIFT [6] and LRP [1], every cell-level patch that contains a certain proportion of positive pixel importance scores is treated as a positive patch and marked in green. The visualization results illustrate that proposed LoopNet can retrieve most of the cancerous cells and avoid most of the healthy areas with only image-level category supervision, hence providing credible evidence of the image classification results. The other approach can only find a few cancerous cells, and plenty of these cells are omitted, so it can not be treated as reliable evidence for clinical diagnosis.

C. Performance on Pseudo Mask Iteration

Like the semi-supervised approach, we can use a trained LoopNet to generate cell locating pseudo masks of the training dataset, then use these generated pseudo masks for computing the cell locating loss $\mathcal{L}_{loc}$, training the LoopNet iter-
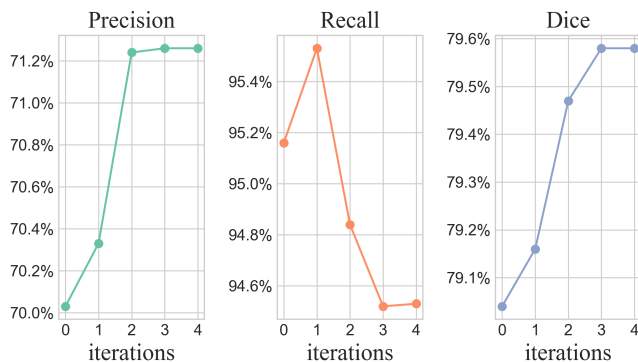

Figure 4. The variation of the metrics along cell locating pseudo mask iterations.

atively. The Fig.4 shows the variation of the metrics along iterations. The "0" iteration means training the LoopNet with original pseudo masks generated by correlation filters. The Precision and Dice increase with the number of iterations, but the Recall increases first and then decreases. This may be because, along with iterative training, the LoopNet can locate cancerous cells more precisely, while some ambigu-

CVPR
#9472

CVPR 2023 Submission #9472. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

CVPR
#9472

ous areas that may contain small parts of the cancerous cells are ignored.

References

[1] S. Bach, A. Binder, G. Montavon, F. Klauschen, KR Müller, and W. Samek. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLOS ONE*, 10, 2015. 2

[2] David S Bolme, J Ross Beveridge, Bruce A Draper, and Yui Man Lui. Visual object tracking using adaptive correlation filters. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2544–2550. IEEE, 2010. 1

[3] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893. Ieee, 2005. 1

[4] João F Henriques, Rui Caseiro, Pedro Martins, and Jorge Batista. High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3):583–596, 2014. 1

[5] Rakshit Naidu and Joy Michael. Ss-cam: Smoothed score-cam for sharper visual feature localization. *arXiv preprint arXiv:2006.14255*, 2020. 2

[6] Avanti Shrikumar, Peyton Greenside, and Anshul Kundaje. Learning important features through propagating activation differences. In *International Conference on Machine Learning*, 2017. 2

[7] Joost Van De Weijer, Cordelia Schmid, Jakob Verbeek, and Diane Larlus. Learning color names for real-world applications. *IEEE Transactions on Image Processing*, 18(7):1512–1523, 2009. 1