AUTOMATIC CLASSIFICATION ON PATIENT-LEVEL BREAST CANCER METASTASES

Sanghun Lee, Joonyoung Cho, Sun Woo Kim

Deep Bio Inc.

ABSTRACT

Automatic diagnosis of breast cancer is a challenge that promises more accessible healthcare. In this paper, we describe the process of predicting slide-level cancer metastasis with machine learning techniques. First, a whole slide image is split into smaller patches which are classified in pixel-level based on Deeplab v3+. Pixel-level classifiers are trained under auto hardmining process. Next, based on pixel-level results for whole slide images, slide-level heatmaps are generated and classified by rule-based criteria for negative, itc, micro and macro. Finally patient-level pN stages are determined by each individual slide-level predictions.

***Index Terms*—** Breast lymph nodes, pN stage, deep neural network, deeplab v3+, hard example mining

1. METHODS

The overall procedure, as illustrated in Figure 1, is as follows:

- Train patch-level classifier with a deep learning model based on Deeplabv3+ under auto hard mining process.
- Predict slides based on maximum lengths from individual heatmaps clustered with DBSCAN algorithm. [1]

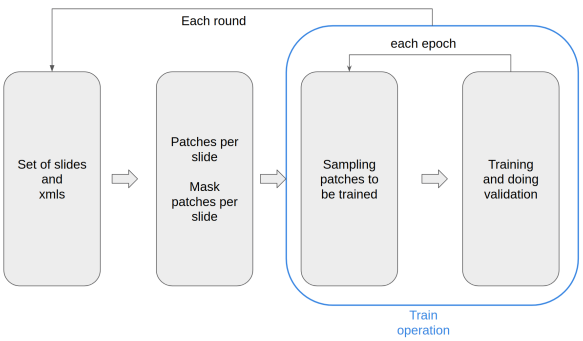

Fig. 1. auto hard mining process

1.1. Data set

The data set for the training patch classifier comes from both Camelyon16 and Camelyon17 dataset (abbreviated to the '16 and the '17 dataset, respectively) [2, 3]. Assuming that the '16 dataset contains two different medical centers, total seven different stain styles are included. Cancerous and non-cancerous patches in '17 dataset are extracted from the slides containing annotations. Training proceeds under the 3-fold cross validation to maximize the utilization of limited number of slides. Total number of slides in the cross validation are 681. For the patient level kappa score, 100 slides except slides from cross validation are tested. Auto hard mining processes are applied to each cross validation set.

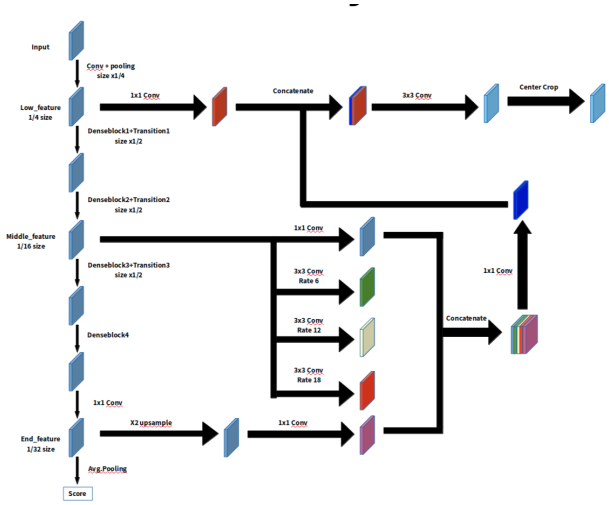

Fig. 2. modified Deeplabv3+

1.2. Train patch-level classifier under auto hard mining process

The patch classifier to predict in pixel level is modified Deeplabv3+ [4] as Figure in 2. The initial learning rate is 0.01 and is reduced by one tenth per 5 epochs. The momentum is randomly chosen out of 0.5, 0.9, 0.99. For each epoch, model inferences whole patches of slides and chooses patches as train set whose intersections over union (IOU) with annotated masks are less than 0.95. This process is

called auto hard mining process. One train process consists of 15 epochs. Total three training process is executed to discover the best model parameters. Between the train processes, the best weight parameter in the predecessor process is set as initial parameter in the next train process. In order to converge faster, initial optimizer is Adam at the first training process and changes to SGD from the second training process. The top three weight parameters ranked from the each cross-validation set are selected to ensemble 3 pixel-level patch classifiers.

1.3. Augmentations

To handle with the micron per meter (mpp) problems, each input patch is re-sized to 704 by 704 whose the original size is larger or smaller within twenty percent compared to the 704 by 704 size. Also, color jitter augmentation is randomly applied with parameters containing brightness, contrast, saturation and hue to overcome color variation. Flip and rotation augmentation are also applied randomly to predict robustly.

1.4. Extracting morphological features

To classify slide-level metastases, morphological features from heatmap are extracted by DBSCAN algorithm. Per slide, major axis is extracted to predict slide. The largest tumor length is the length of major axis in the pixel unit times mpp.

2. RESULTS

Mean IOU of pixel-level classifiers is 0.901 from validation slides. For the 100 test slides, the mean slide-level accuracy of the three ensembled models is 0.93, and mean kappa score is 0.93.

3. DISCUSSION

The previous approaches of training patch classifier are based on fixed patch samples. That method shows restricted patch sampling efficiency in the case of limited train dataset. Under auto hard mining process, the patch classifier can be trained with effective sampled patches.

To handle with itc cases which are so small to detect so that it is hard to distinguish whether these are false positives or not, patch classifier is chosen as a segmentation model. Based on our experiments, segmentation models shows less false positive after post processes.

In the field of pathology, it costs too much to get a area annotated data. The semi-supervised learning with only slide diagnosis is good option to tackle this problem, but still that method requires the huge number of dataset, at least over 10 thousands slides.

4. REFERENCES

[1] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu, "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. 1996, KDD'96, pp. 226–231, AAAI Press.

[2] Babak Ehteshami Bejnordi and Jeroen van der Laak, "Camelyon16: Grand challenge on cancer metastasis detection in lymph nodes," 2016.

[3] Oscar Geessink, Péter Bándi, Geert Litjens, and Jeroen van der Laak, "Camelyon17: Grand challenge on cancer metastasis detection and classification in lymph nodes," 2017.

[4] George Papandreou Florian Schroff Hartwig Adam Liang-Chieh Chen, Yukun Zhu, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *ECCV 2018*, 2018.