BREAST CANCER PN-STAGE CLASSIFICATION FOR WHOLE SLIDE IMAGES

Zixu Zhao, Huangjing Lin, Pheng Ann Heng

Department of Computer Science and Engineering, The Chinese University of Hong Kong

ABSTRACT

Pathologic N-stage (pN-stage) classification is an essential task for breast cancer patient's diagnosis and treatment. This challenging task requires accurate prediction on whether the cancer has metastasized to the regional lymph nodes in whole slide histological images. In this paper, we propose a novel framework to automatically predict pN-stage for whole slide images using our multi-scale ScanNet. Our model is evaluated on Camelyon 16 and Camelyon17 datasets.

Index Terms— Camelyon17, Deep learning, pN-stage Classification, Breast Cancer

1. METHOD

The basic algorithmic steps of our method are: preprocessing, slide-level metastasis detection, feature extraction, and patient-level classification. We start with identifying tissue regions and removing non-informative regions on the WSIs. Subsequently, we feed image tiles extracted from the tissue regions into our multi-scale ScanNet for efficient and dense prediction at slide level. After that, we threshold the metastasis-likelihood maps we obtained and extract several features from the predicted cancerous area. We finally train a random forest classifier with these features and then predict the pN-stage of each patient.

1.1. Preprocessing

To filter the regions without tissue and determine the adaptive threshold, we employ the OTSU algorithm [1] at a low resolution level (level 4). The thresholds are applied in RGB (red-green-blue) color space. No more morphological operations are used to refine thresholded images.

1.2. Slide-level Detection

We propose a new version of ScanNet [2] to detect the cancer metastasis in a multi-scale manner. To avoid the boundary effect of FCN predictions, the paddding operations are removed which is the same as ScanNet. We increase the input resolution from a size of 244×244 to a size of 436×436 to capture more detailed information which is hard to learn. The last two convolutional layers (*Conv*6 and *Conv*7) of ScanNet are replaced with a dense module, in which the extracted

low-level features and high-level features are independently cropped before average pooling layers. Here, we control the pooling stride with a dense coefficient for dense reconstruction. Finally, the concatenated features are fed into the last convolutional layer for classification. Figure 1 shows the basic architecture of our multi-scale ScanNet.

In the training phase, small image tiles are randomly cropped from WSIs with a size of 436×436 to train the multi-scale ScanNet. To increase the variation in the training set, we perform extensive data augmentation including random flipping, scaling, rotation over angles between 0 and 360, and cropping. In addition, we perform color jittering and HSV augmentation to make our model robust to color variation caused by differences between scanners and staining protocols. Our models are trained with hard negative mining.
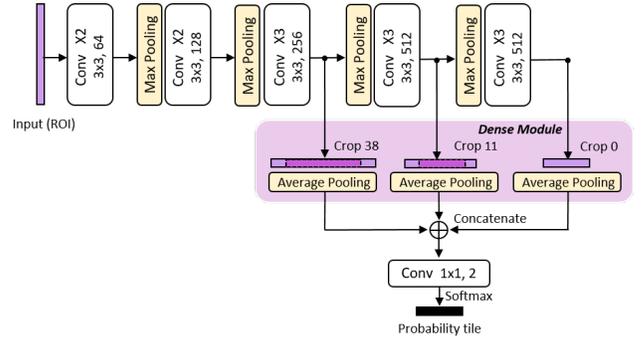

Fig. 1. The architecture of multi-scale ScanNet

In the prediction phase, by leveraging the merit of fully convolutional architecture, our model can take a region of interest(ROI) with a size up to 2484×2484 (determined by the maximal capacity of GPU memory) as input and output a probability tile with a size of 72×72. In this way, our model can process a WSI hundreds of times faster than patch-based classification methods with the same stride. In addition, we can densely and accurately reconstruct the whole-slide probability map by using the dense module in our model.

1.3. Feature Extraction

Each whole-slide probability map is converted into a feature vector after being thresholded with 5 threshold value, i.e., 0.5, 0.6, 0.7, 0.8, 0.9, which is then used to train a lymph node

classifier. According to the morphological and geometrical information, we extract 13 types of features from each thresholded probability map (65 features in total).

1.4. Patient-level Classification

We use a random forest classifier to classify lymph nodes into four classes (Normal, Isolated tumor cells (ITC), Micro, Macro). All the feature vectors are fed into the the random forest classifier for slide-based training and prediction. Finally, the patient's pN-stage is graded according to the given rules.

2. EXPERIMENTS

We evaluated our method on the benchmark dataset of Camelyon16 [3] and Camelyon17 [4] challenges. For training multi-scale ScanNet, we choose 400 WSIs from Camelyon16 dataset (including 208 slides with annotations) and 215 WSIs of 43 patients from Camelyon17 dataset (including 50 slides with annotations) as the training data (we name it train-*a* set). For training the random forest classifier, we use the remaining slides from Camelyon17 dataset (including 55 abnormal slides without annotations and 285 WSIs from the remaining 57 patients) as the training data (we name it train-*b* set).

In this submission, we generate our results on the testing data from Camelyon17 via a well-trained single model. Before submission, we evaluate this model on the train-*b* set with by five-fold cross validation. We split the train-*b* set into five subsets and randomly choose one subset for validation and the others for training the random forest classifier. We adopt the quadratic weighted kappa score as the evaluation metric. The patient-level accuracy and slide-level accuracy of our model are 0.9266 and 0.9363.

3. REFERENCES

[1] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan 1979.

[2] Huangjing Lin, Hao Chen, Qi Dou, Liansheng Wang, Jing Qin, and Pheng-Ann Heng, "Scannet: A fast and dense scanning framework for metastatic breast cancer detection from whole-slide images," 03 2018.

[3] "Camelyon 2016," http://camelyon16.grand-challenge.org/, 2017, [Online; accessed 24-July-2017].

[4] "Camelyon 2017," http://camelyon17.grand-challenge.org/, 2017, [Online; accessed 24-July-2017].