PN-STAGE CLASSIFICATION WITH DEEP LEARNING

Huangjing Lin [†,§] *Hao Chen*[†,§] *Qi Dou*[†,§] *Liansheng Wang*[‡] *Pheng-Ann Heng*[†]

[†] Imsight Medical Technology, China
[§] Department of Computer Science and Engineering, The Chinese University of Hong Kong
[‡] Department of Computer Science, Xiamen University, China

1. METHOD

We first employ a simple yet efficient method to remove the non-informative regions of input WSI. Then we feed pre-processed images into the modified FCN equipped with reconstruction algorithm for efficient and dense predictions. After that, we extract the features from the predictions, and classify the whole-slide images via a random forest classifier. Finally, the patient pN-stages finally are generated based on the staging rules and WSI staging information.

1.1. Pre-processing

It is observed that more than 70% area of a typical WSI is covered by the non-informative background that provides almost no information for cancer assessment. In order to remove these regions to save computational cost, we employ the simple OTSU algorithm [1] to determine the adaptive threshold and filter out most of the white background.

1.2. Fast and Prediction via FCN

We propose to harness a modified fully convolutional network for fast prediction for large WSIs by taking its advantage of allowing us to take arbitrary sized images as input (Architecture as Table. 1). Different from traditional FCNs that are commonly used for segmentation tasks [2], our FCN has no upsampling path which is a must for segmentation but not necessary for detection tasks. We implement the proposed FCN based on a modified VGG-16 [3] network by removing padding operations and replacing the last three fully connected layers with fully convolutional layers $1024 \times 1024 \times 2$ (i.e., kernel size 1×1). Thus, our FCN can enjoy the transferred features learned from a large set of natural images [2]. In the training process, we employ patch samples with size as 244×244 randomly cropped from WSIs to train the FCN.

1.3. Reconstruction for Dense Prediction

Reconstruction of the whole-slide probability map from the set of probability tiles generated by the FCN for accurate dense predictions is one of the key steps of the proposed ScanNet and it is not trivial. As the set of probability tiles are generated from a set of sub-images extracted from the input WSI by a certain offset, we call the set of probability tiles as *offset probability tiles* (OPTs). We propose a two-stage scheme to reconstruct the whole-slide probability map. We first generate a set of *dense probability tiles* (DPTs) based on the OPTs and then stitch these DPTs together to obtain the final probability map.

Table 1. The Architecture of ScanNet.

Layer	Feature maps(Train)	Kernel size	Stride
Input	$244 \times 244 \times 3$	-	-
Conv1_1	$242 \times 242 \times 64$	3×3	1×1
Conv1_2	$240 \times 240 \times 64$	3×3	1×1
Pool1	$120 \times 120 \times 64$	2×2	2×2
Conv2_1	$118 \times 118 \times 128$	3×3	1×1
Conv2_2	$116 \times 116 \times 128$	3×3	1×1
Pool2	$58 \times 58 \times 128$	2×2	2×2
Conv3_1	$56 \times 56 \times 256$	3×3	1×1
Conv3_2	$54 \times 54 \times 256$	3×3	1×1
Conv3_3	$52 \times 52 \times 256$	3×3	1×1
Pool3	$26 \times 26 \times 256$	2×2	2×2
Conv4_1	$24 \times 24 \times 512$	3×3	1×1
Conv4_2	$22 \times 22 \times 512$	3×3	1×1
Conv4_3	$20 \times 20 \times 512$	3×3	1×1
Pool4	$10 \times 10 \times 512$	2×2	2×2
Conv5_1	$8 \times 8 \times 512$	3×3	1×1
Conv5_2	$6 \times 6 \times 512$	3×3	1×1
Conv5_3	$4 \times 4 \times 512$	3×3	1×1
Pool5	$2 \times 2 \times 512$	2×2	2×2
Conv6	1024	2×2	1×1
Conv7	1024	1×1	1×1
Conv8	2	1×1	1×1

1.4. pN-Stage Prediction by Random Forest

Finally, we use a random forest classifier to grade the stages of WSIs. Our derived features are based on a previous work of [4] by thresholding the predictions with five pre-defined values: $0.5, 0.6, 0.7, 0.8, 0.9$. In each binary mask, we extract the max probability, mean probability, max axis length, positive percentage in tissue, area and convex area from its highest prediction region and largest area region respectively (60 features in total). The features are fed to random forest classifier for slide-based training and prediction. Finally, the patient-based staging is conducted based on the slide-based staging.

2. EXPERIMENTS AND DATASET

We validated our method on the benchmark dataset of *Camelyon16* and *Camelyon17* challenge [5]. The training data contained 350 normal WSIs (240 WSIs from *Camelyon16* and 110 WSIs from *Camelyon17*) and 208 tumor WSIs (158 WSIs from *Camelyon16* and 50 WSIs from *Camelyon17*) with pixel-level annotations.

Submission2
In this submission, we generated our results on the testing data of Camelyon17 via averaging results from two trained models.

3. REFERENCES

[1] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[2] Hao Chen, Xiaojuan Qi, Lequan Yu, Qi Dou, Jing Qin, and Pheng-Ann Heng, "Dcan: Deep contour-aware networks for object instance segmentation from histology images," *Medical Image Analysis*, vol. 36, pp. 135–146, 2017.

[3] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[4] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H Beck, "Deep learning for identifying metastatic breast cancer," *arXiv preprint arXiv:1606.05718*, 2016.

[5] "2016 ISBI Challenge on Cancer Metastasis Detection in Lymph Node," https://camelyon16.grand-challenge.org/home/.