

AI Breast Cancer Tumor Detection

Manhattan

ABSTRACT

Predicting TNM stage is the major determinant of breast cancer patient's prognosis and treatment. The essential part of TNM stage classification is whether the cancer has metastasized to the regional lymph nodes (N-stage). Pathologic N-stage (pN-stage) is commonly performed by pathologists detecting metastasis in histological slides. However, this diagnostic procedure is prone to misinterpretation and timeconsuming. Automated detection of lymph node metastasis and pN-stage prediction has a great potential to reduce their workload and help the pathologist. In this paper, we present a framework to automatically predict pN-stage from whole slide histopathology images. pN-stage is predicted by combining convolutional neural network (CNN) based metastasis detector and slide-level lymph node classifier module. Our framework is evaluated on Camelyon17 which is recently introduced challenging benchmark dataset. Index Terms— Camelyon17, Convolutional neural network, Deep learning, Metastasis detection

1. INTRODUCTION

The TNM stage system [1] is widely used to classify the magnitude of cancer spread which is a significant component of breast cancer control and surveillance. The essential part of TNM stage classification is whether the breast cancer has spread to the regional lymph nodes (N-stage) since lymph nodes are the first place breast cancer is likely to metastasize. N-stage is commonly determined by metastasis detection which is performed in lymph node histological slides.

However, the diagnostic procedure examined by pathologists to detect metastases is prone to misinterpretation and timeconsuming and tedious. Automated detection of lymph node metastasis and pN-stage prediction has a great potential to reduce their workload and help the pathologist. In the last few years, considerable improvements have been emerged in the computer vision task using convolutional neural network (CNN) [2]. Followed by this paradigm, CNN based computer assisted metastasis detection has been proposed in recent years [3, 4]. In [3], the author proposed the unified framework for tumor proliferation score prediction in breast histopathology which was used to win the Tumor Proliferation Assessment Challenge at MICCAI 2016 [5]. [4] suggested CNN based lymph node breast tumor detection framework which obtained state-of-the-art results on the Camelyon16 [6] dataset. In this paper, we propose an automatic framework to predict pathologic N-stage (pN-stage) from patient's whole slide histopathology images. The proposed framework is conducted by integrating three modules: a region of interest (ROI) extraction module, a CNN-based metastasis detection module, and a slide-level lymph node classification module. First, ROI extraction module proposes candidate tissue regions from whole slide images. Second, CNN-based metastasis detection module predicts cancer metastasis within extracted ROIs. Third, the predicted scores extracted from ROI are converted to a feature vector based on the morphological and geometrical information which is used to build a slidelevel lymph node classifier. Finally, patient-level pN-stage is determined by aggregating slide-level predictions.

2. METHOD

We have trained a ResNet-50 network with Camalyon'16 and Camleyon'17 datasets. We used 50 annotated tumor slides and 70% of negative slides of Camleyon'17 for training and remaining negative slides for validation.

2.1 Patch Extraction

A whole slide image (WSI) is approximately 200000 x 100000 pixels on the highest resolution level. If we deal with all regions, enormous computation time is required because of the huge size of the slide. In order to extract tissue regions from the WSIs, Otsu threshold [7] or gray value threshold [8] is commonly used in recent studies. We observed that metastasis regions are commonly located at the edge of the tissue regions. Therefore, careful tissue region extraction method is needed. We determine to use gray value threshold method which is obtained a metastasis region's sensitivity 0.9752 on Camelyon16 train set. In detail, we convert RGB to gray from 32-times down-sampled WSI and then extract tissue regions by thresholding gray value > 0.8 [8].

2.2 Data Augmentation

We have use very diverse set of data augmentation to incorporate all possibilities. This includes- Random Image Rotation, Random Image left right flip, Random Image Up and Down Flip, Random Hue, Random Saturation and Random Brightness. Sample of augmented images with rotation and HSV color variations Figure-1.

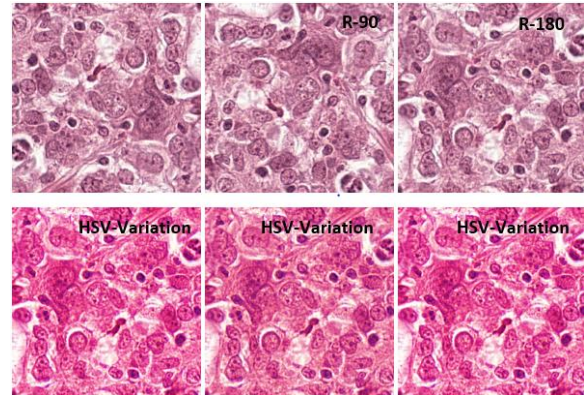


Figure -1 Data Augmented Samples

2.3 Convolutional Neural Network

Some annotated metastasis regions include fat tissues and non-metastasis area since accurate pixel-level annotation is difficult in large size WSIs. Deep learning model becomes robust to noisy labels when a larger dataset is available [9]. Therefore, we build a large scale dataset by extracting small patches from WSIs to deal with those noisy labels. After the ROIs are founded from WSIs, we extract 256x256 patches within ROIs with stride 128. We label a patch as tumor if over 75% pixels in the patch are annotated as a tumor. Our metastasis detection module is based on the wellknown CNN architecture ResNet101 [2] for patch classification to discriminate between tumor and non-tumor patches. Training CNN model with extracted patches from WSIs is challenging because a number of extracted patches is various per WSI. To deal with this imbalance, we followed similar patch sampling approach used in [4]. In detail, we sample the same number of tumor/normal patches where patches are sampled from each slide with uniform distribution. To combat with the variety of hematoxylin and eosin (H&E) stained color because of chemical preparation difference per slide, extensive color augmentation is performed by applying random hue, saturation, brightness, and

contrast. Since the classes of histopathology image exhibit rotational symmetry, we include data augmentation by randomly rotating over angles between 0 and 360, and random left-right flipping.

2.4 Heatmap Generation

For each whole slide image, after extracting the tissue region from a lower magnification, we take the corresponding patch at highest magnification and run through our trained model, using the obtained probability we generate a probability heatmap for the whole slide image.

3. Whole Slide Classification

To determine each patient’s pN-stage, lymph node should be classified into four classes (Normal, Isolated tumor cells (ITC), Micro, Macro). For each lymph node WSI, we obtain the 128-times down-sampled tumor probability heatmap through the CNN based patch classifier described beforehand. Each heatmap is converted into a feature vector which is used to build a slide level lymph node classifier. We define 11 types of features based on the morphological and geometrical information. By using converted features, XGBoost classifier is trained to automatically classify the lymph node into four classes. Finally, each patient’s pN-stage is determined by aggregating all lymph node predictions with the given rule.

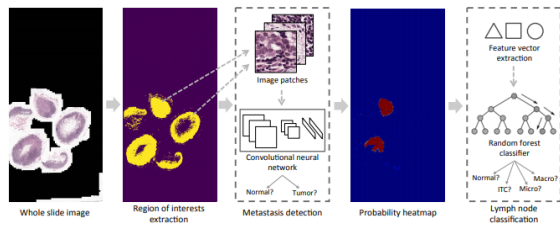


Figure -2 Algorithm Pipeline

heatmaps. Some of those features are listed below in the table-

Max probability in biggest tumor blob	Maximum probability in whole slide tumor
Avg probability in biggest tumor blob	Area of ellipse fit on biggest tumor blob
Major axis of ellipse fit of biggest tumor blob	Solidity of ellipse fit on biggest tumor blob
Minor axis of ellipse fit on biggest tumor blob	
Equivalent Diameter of biggest tumor blob	

We have extract 15 features for each threshold of heatmaps. We have used two threshold values.

XGboost model is trained with 5-fold cross validation.

4. RESULTS

We have trained our model first only on CAMELYON16 Training dataset, to evaluate how good is it. We achieved a FROC score of 0.65.

Using the same model, we trained it on the remaining datasets (Camelyon16 Test as well as Camelyon17 Training). Our whole slide classifier was trained using the features extracted from the heatmaps generated by above model. Our results are as follows-, our five-fold cross validation average scores are as follows.

Model	Score	Metric
ResNet-50	0.65	FROC

XGboost (overall)	0.89(Train) 0.74(Avg 5- folds Cross Validation)	Quadratic Kappa
----------------------	--	--------------------

We have achieved competitive performance on our tumor classification model and whole slide image classification model.

4. References

[1] Meier, Frederick A. "The Landscape of Error in Surgical Pathology." *Error Reduction and Prevention in Surgical Pathology*. Springer New York, 2015. 3-26.

[2]https://camelyon16.grand-challenge.org/site/CAMELYON16/serve/public_html/Results/Presentations/Camelyon16_BIDMC_CSAIL.pdf/

[3] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979

[4] Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q Nelson, Greg S Corrado, et al., "Detecting cancer metastases on gigapixel pathology images," *arXiv preprint arXiv:1703.02442*, 2017.

[5] Andrew Janowczyk and Anant Madabhushi, "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases," *Journal of pathology informatics*, vol. 7, 2016.

[6] David Rolnick, Andreas Veit, Serge Belongie, and Nir Shavit, "Deep learning is robust to massive label noise," *arXiv preprint arXiv:1705.10694*, 2017.

[7] "Camelyon 2017," <https://camelyon17.grand-challenge.org/>, Accessed: 2017-07-24.

[8] Mart'ın Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al., "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.