ACCURATE AND FAST PREDICTION OF BREAST CANCER METASTASIS STAGES VIA INCEPTION BASED SEMANTIC SEGMENTATION NETWORKS

Shenghua Cheng, Shaoqun Zeng, Jiangsheng Yu

Huazhong University of Science and Technology, Bingsheng Technology Company, Wuhan, China

ABSTRACT

Automatic and accurate detection of breast cancer metastases is meaningful for reducing the workload of the pathologists and reducing the cost of diagnosis. The key of predicting breast cancer metastasis stages is to how to get probability heatmap of tumor region prediction accurately and fast. The common CNN-based scanning strategy at a stride is a tradeoff between speed and accuracy. In this paper, we proposed an inception structure based semantic segmentation network for accurate and fast prediction of breast cancer regions. This method has several advantages: 1) high prediction speed derived from dense pixel classification of the semantic segmentation network; 2) high accuracy stem from level 0 heatmap, division of WSI with sufficient overlapping regions and big patch size of more rich content. This method is trained and validated on Camelyon17 challenge datasets.

Index Terms – Cancer metastases prediction, deep convolutional networks, Camelyon17 challenge

1. INTRODUCTION

Breast cancers metastasis diagnosis needs the analysis of hematoxylin and eosin stained slides [1]. Currently, the slides are examined done by pathologists manually. But this manual examination depends on the experience of the pathologists and is subjective and time-consuming. The centimeter-scale sections are imaged at sub-micrometer resolution, thus, the resulting whole slide images (WSIs) are of Gigapixels. Manual check of the WSIs is very tedious and easily misses small regions. Therefore, automatic and accurate detection of breast cancer metastases is meaningful for reducing the workload of the pathologists and reducing the cost of diagnosis. Recent Camelyon16 and 17 challenges promote the development of cancer automatic detection algorithms in pathology WSIs [2-4]. The key of these methods is how to predict tumor regions in WSIs at pixel level. Compared traditional model-based or hand-crafted features methods, deep learning is data-driven methods and can learn hierarchical feature representation automatically [5]. In image classification and semantic segmentation fields, deep convolutional networks have archived great successes in recent years [6-7]. The common CNN-based scanning strategy at a stride is a tradeoff between speed and accuracy. In this paper, we proposed an inception structure based semantic segmentation network for accurate and fast prediction of breast cancer regions.

The goal of Camelyon17 [8] is to predict the pN-stage of each patient, which is decided by the metastasis status of his five lymph node WSIs. Thus, the key is to accurately predict WSI metastasis status (normal, isolated tumor cells, micro-metastases and macro-metastases). To solve this problem, this paper presented a pipeline consisting of three parts (**Fig. 1**): tissue extraction, tumor region detection and patient pN-stage prediction. The tissue region extraction is based on the difference of pixel RGB values. The tumor regions in WSIs are detected by the proposed semantic segmentation network based on GoogLeNet-v3 finetuned on the ImageNet weights [7]. The WSI status is classified by random forest algorithm with morphological features extracted from the predicted tumor heatmap, then the pN-stage can be obtained by the given rule. There are three key points for accurately predicting cancer metastasis stages: 1) effectively deal with the imbalance of tumor and normal samples and the imbalance of tumor regions of various cases (isolated tumor cells, micro-metastases and macro-metastases); 2) how to train the very deep network on the limited datasets of simple samples and hard samples; 3) how to prevent overfitting when learning a classifier of WSI metastasis status on 500 WSIs. Aiming at these challenges, we designed some effective strategies for training sample preparation (Section 2.3.1), network training (Section 2.3.2) and the WSIs metastasis status classification (Section 2.4).

2. METHODS

2.1. Datasets

The training and testing datasets come from Camelyon16 [9] and 17 challenges [8] . The finished Camelyon16 challenge contains 160 tumor WSIs with tumor region labelling and 240 normal WSIs. The Camelyon17 challenge contains 100 patients (5 WSIs per patient) with pN-stage classification. 50 WSIs of the 500 WSIs have tumor region annotation. Another 100 patients are provided for testing. In tumor region detection stage, we used all the WSIs with annotation and normal WSIs as our training datasets.

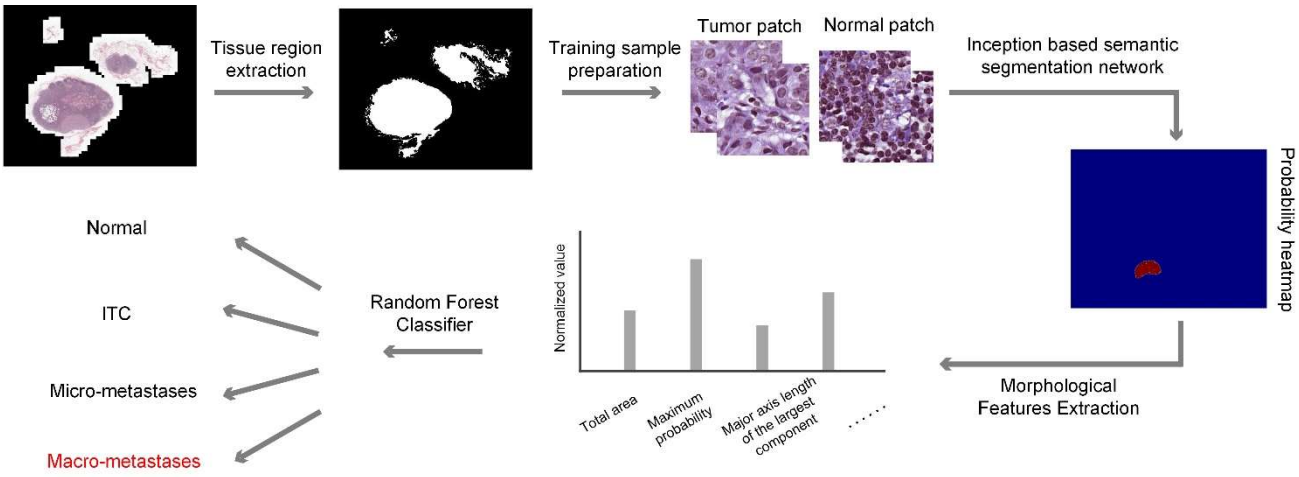

Fig. 1. The diagram of the presented learning-based method of cancer pN-stage prediction.

2.2. Extraction of Tissue Region

The whole slide pathological images (WSIs) usually contain a lot of non-tissue blank regions. Thus, we design a method to extract the tissue regions from the WSIs . Post tumor region detection is then constrained on the extracted tissue regions. After observation, we found that the tissue and non-tissue regions can be classified by the pixel color. The colored region are more possibly be the tissue region. Therefore, we use the different of the maximum and minimum values of pixel RGB values to evaluate the colored degree. Tissue foreground masks can be obtained with a threshold 10. Holes on the masks are filled with morphological operations. We found this method is simpler than the common OTSU method [10] but is valid for tissue region extraction in WSIs (**Fig. 2**).

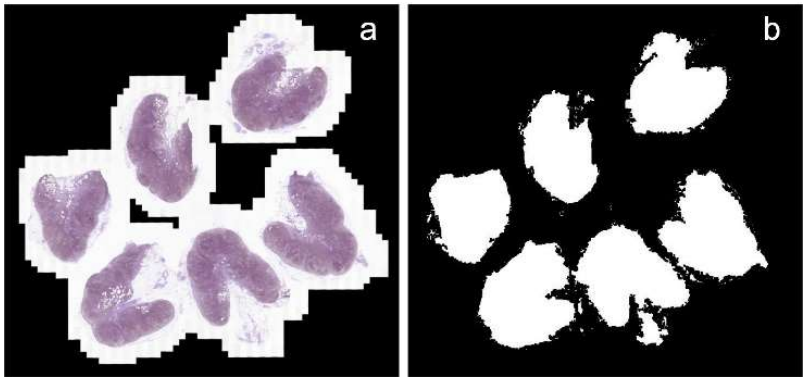

Fig. 2. RGB-based tissue region extraction. **a** is the original WSI of patient 0 node 2. **b** is the extracted tissue mask. If the different of the maximum and minimum values of a pixel RGB values is greater than 10, the pixel is labeled as tissue region. Holes on the masks are filled with morphological operations. All tissue regions in the WSI have been labelled in its entirety using the simple method.

2.3. Tumor Region Detection by Inception Based Semantic Segmentation Network

We proposed an inception structure based semantic segmentation network for accurate and fast prediction of breast cancer regions. This method has several advantages: 1) high prediction speed derived from dense pixel classification of the semantic segmentation network; 2) high accuracy stem from level 0 heatmap, division of WSI with sufficient overlapping regions and big patch size of more rich content. The detailed structure of the network is depicted in **Fig. 3.**

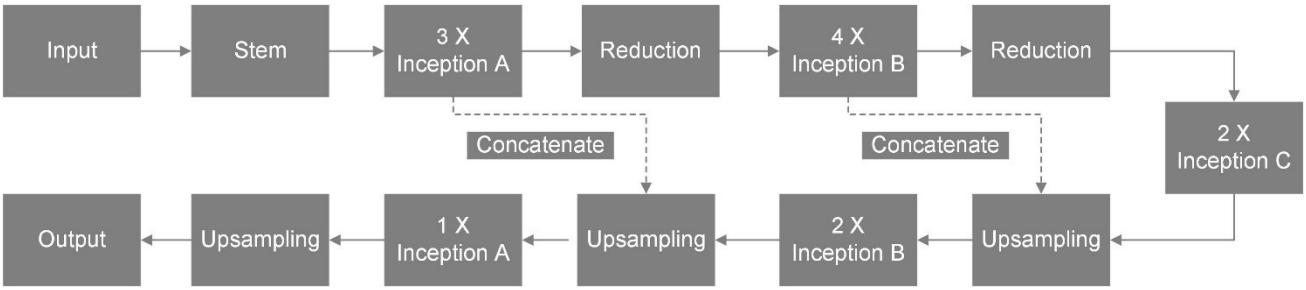

Fig. 3. The network architecture of the proposed semantic segmentation network based on inception structures.

2.3.1. Preparation of training samples

From each tumor WSIs with annotation, we randomly (with uniform distribution) select 1K positive patches in tumor regions and 10K negative patches in normal regions. From each normal WSIs, we randomly select 10K negative patches. For each training epoch, we randomly sampling the same number of positive/negative patches. The strategy has several advantages: 1) positive patches are sampled from each tumor WSIs with the same probability, thus, WSIs with small tumor regions will be not be ignored in the training; 2) the imbalance of positive and negative patches is dealt with by resampling. The patch size is set to 512×512, which can contain enough cells in a patch as recognition context. The RGB values of a patch are normalized into [-1 1]. No sample augmentation has been done.

2.3.2. Network training

The proposed semantic segmentation network was trained three rounds. First we trained the corresponding CNN of the semantic segmentation net to classify normal patch and tumor patch. Then we utilized the learned weights of the CNN and only trained the rest weights of the semantic segmentation net. Finally, we found the mislabeled pixels and extracted neighboring regions as hard samples for the third round training. The step-by-step finetuning strategy has the following advantages: 1) we take full use of the weights of the corresponding CNN; 2) the network is easier to achieve convergence from the shallow net to the deep net and from the simple samples to hard samples. The network is trained with Adam method [11].

The heatmap generated by the proposed net is more accurate and with a higher resolution than the heatmap generated by the common CNN-based scanning method. The CNN-based method is a tradeoff between speed and accuracy, which usually has to use a large stride for speed. Thus, CNN heatmap is coarse and has much noise especially in the small tumor regions. The heatmap of our method is generated at level 0 with a fast speed. In addition, we utilized a patch size of 512 pixels and an overlapping width of 128 pixels to ensure big recognition content and remove boundary effect.

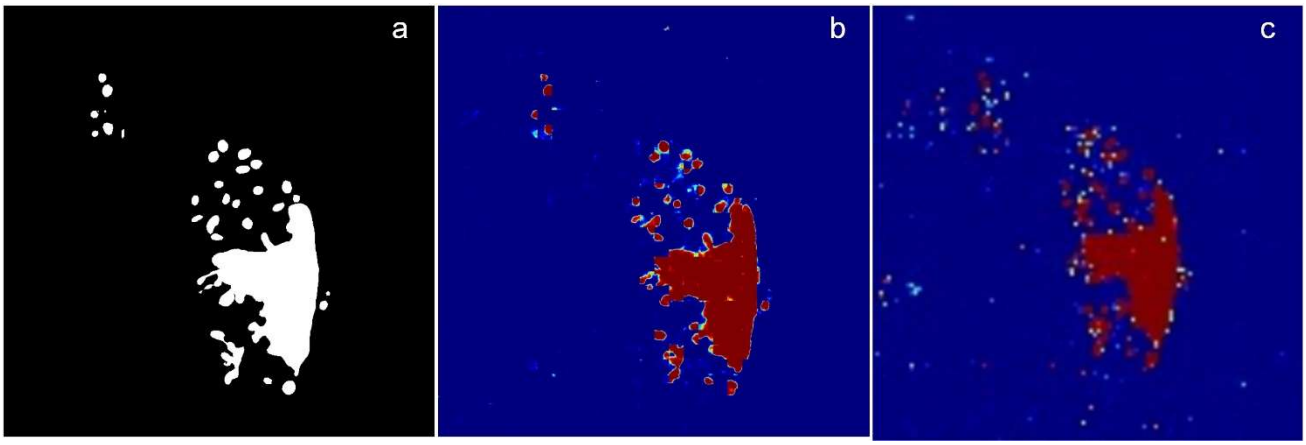

Fig. 4. Comparison in prediction heatmap between the CNN-based scanning method and the proposed semantic segmentation net. **a** Ground truth of the annotated tumor regions. **b** The heatmap of the proposed method. **c** The heatmap of the CNN-based scanning method with stride 128. The heatmap of the proposed method is more accurate and with a higher resolution.

2.4. Prediction of Patient pN-stage

For each WSI, we divide it into a number of patches with sufficient overlapping regions. The patch size is 512 x 512 and the overlapping width is 128 pixels. Then, we can obtain a tumor region prediction heatmap of each WSI at level 0 using the trained network. Some morphological operations are used to fill the holes and close the small gaps of the binarized heatmap. 13 morphological features (**Table 1**) of a binarized WSI heatmap are extracted for classification the WSI into four states: normal, ITC (isolated tumor cells), micro-metastases and macro-metastases. Since we only have 500 WSIs for training and validation, the key of this part is how to prevent overfitting with the small dataset. We use random forest algorithm to learn WSI state classification on the Camelyon17 datasets. K-fold cross validation is adopted for preventing overfitting and adjusting hyperparameters (WSI heatmap binarization threshold and random forest hyperparameters). The randomness of feature selection and subset selection when generating decision trees helps preventing overfitting. Finally, we can obtain the patient pN-stage using the given rule.

Table 1. The 13 morphological feature of a binarized WSI heat map.

Feature Number	Feature Name	Feature Number	Feature Name
1	Total area	8	Maximum probability of the largest component
2	Maximum probability	9	Minor axis length of the largest component
3	Average probability	10	Average probability of the largest component
4	Area of the largest component	11	Extent of the largest component
5	Convex area of the largest component	12	Solidity of the largest component
6	Major axis length of the largest component	13	Eccentricity of the largest component
7	Equivalent Diameter of the largest component		

3. EXPERIMENTS

We trained our tumor region detection network on Camelyon16 and 17 datasets (all manually labelled tumor WSIs and all normal WSIs). The dice score of training achieved 0.98 within one weeks. The training was implemented in Keras [12] with Tensorflow [13] backend using two Titan XP. In application stage, we got the heatmaps with a patch size of 512 pixels and an overlapping width of 128 pixels to ensure big recognition content and remove boundary effect. Generating one WSI heatmap needs about 5~10 minutes using two Titan XP.

The random forest for classifying the WSI status was trained on Camelyon17 training dataset (500 WSIs). Because the small number of cases, 5-fold cross validation was utilized to search good hyperparameters. We obtained 0.88 average validation accuracy of WSIs status classification.

4. REFERENCES

[1] Fischer, A. H. et al. Hematoxylin and eosin staining of tissue and cell sections. *Cold Spring Harbor Protocols* **2008**, pdb-prot4986 (2008).

[2] Litjens, G. et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific Reports* **6**, 26286 (2016).

[3] Wang, D. et al. Deep learning for identifying metastatic breast cancer. *arXiv preprint arXiv*, 1606.05718 (2016).

[4] Liu, Y. et al. Detecting cancer metastases on gigapixel pathology images. *arXiv preprint arXiv*, 1703.02442 (2017).

[5] LeCun, Y., et al. Deep learning. *Nature* **521**, 436-444 (2015).

[6] Krizhevsky, A. et al. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097-1105 (2012).

[7] Szegedy, C. et al. Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2818-2826 (2016).

[8] Geessink, O. et al. Camelyon17: Grand challenge on cancer metastasis detection and classification in lymph nodes, 2017.

[9] Bejnordi, B. E. et al. Camelyon16: Grand challenge on cancer metastasis detection in lymph nodes, 2016.

[10] Otsu, N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* **9**, 62–66 (1979).

[11] Kingma, D. et al. Adam: A method for stochastic optimization. *arXiv preprint arXiv*, 1412.6980 (2014).

[12] Chollet, F. Keras (2015). Available at: http://keras. io (2017).

[13] Bastien, F. et al. Theano: new features and speed improvements. *arXiv preprint arXiv*, 1211.5590 (2012).