

# BREAST CANCER METASTASES DETECTION

## ABSTRACT

Automated detection and classification of breast cancer metastases in whole-slide images of histological lymph node sections is meaningful in the real life for reducing the cost for Cancer Screening and make it possible for everyone to get regular check-up to lessen their risk for cancer. We propose an automated diagnosis framework based on deep learning for breast cancer metastases.

**Index Terms**— Detection, Breast Cancer Metastases , Automated Diagnosis,

## 1. INTRODUCTION

Metastatic breast cancer is breast cancer that has spread beyond the breast to other organs in the body. Typically , cancer cells travel through the walls of nearby lymph vessels. at least 154,000 people in the U.S. have metastatic breast cancer. Some women have metastatic breast cancer when they are first diagnosed. Automated detection and classification of breast cancer metastases is helpful for early regular check-up to treat it in advance.

## 2. DATASET AND EVALUATION METRICS

In this section , we describe the Camelyon17 dataset provided by the Diagnostic Image Analysis Group (DIAG) and Department of Pathology of the Radboud University Medical Center and the evaluation metrics to rank the diagnosis algorithm.

### 2.1. Dataset

The Camelyon17dataset consist of a total 1000 whole slides images(WSI) from 200 patients(1 patient have 5 whole-slide images of histological lymph node sections). It is split into 100 patients for training and 100 patients for testing. The patients' WSIs comes from different center where the approach of staining differs a lot.

### 2.2. Diagnosis target

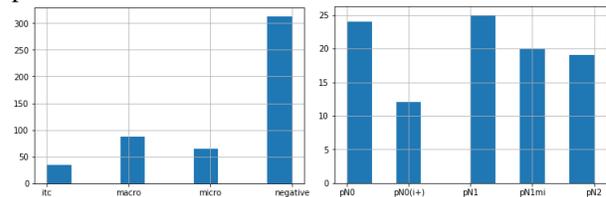
Our diagnosis target is determining a pN-stage for every patient. There are two categories of lymph node metastasis: Macro-metastases: metastases greater than 2.0 mm. Micro-metastases: metastases greater than 0.2 mm or more than 200 cells, but smaller than 2.0 mm. Typically, single tumor cells or a cluster of tumor cells smaller than 0.2 mm or less than 200 cells. Lymph nodes containing only ITC are

therefore not counted as positive lymph nodes. However, pathologists are required to report on ITC when no macro-metastases or micro-metastases were found in a patient's lymph nodes. Based on the description of diagnosis on WSI level, the pN stage on patient level can be defined as below:

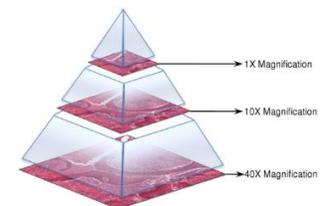
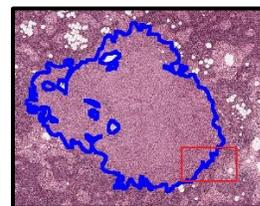
- pN0: No micro-metastases or macro-metastases or ITCs found.
- pN0(i+): Only ITCs found.
- pN1mi: Micro-metastases found, but no macro-metastases found.
- pN1: Metastases found in 1–3 lymph nodes, of which at least one is a macro-metastasis.
- pN2: Metastases found in 4–9 lymph nodes, of which at least one is a macro-metastasis.

### 2.3. Dataset details

The specific label distribution in WSI level and patient level provided in table 1 and table 2.



- Slides were collected from 5 different medical centers for CAMELYON17.
- 3 different scanners were used.
- The slides were anonymized, and converted to standard TIFF file format.
- Image is stored in a multiresolution format in TIFF.
- Pixel size was 0.25×0.25 μm on the lowest level.
- Pixel count: approximately 20 gigapixels.
- 10 slides from each center were exhaustively annotated



### 2.4. Evaluation Metric

Five class quadratic weighted kappa was used for evaluation where the classes were the pN-stages.

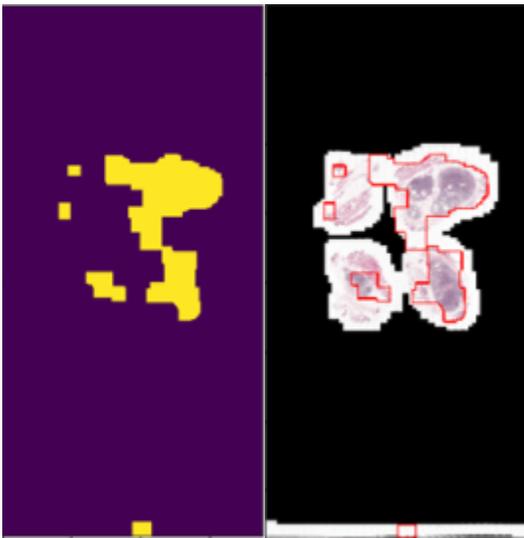
$$\kappa = \frac{\sum_{i=1}^k \sum_{j=1}^k w_{ij} x_{ij}}{\sum_{i=1}^k \sum_{j=1}^k w_{ij} m_{ij}}$$

where k=number of classes and  $w_{ij}$ ,  $x_{ij}$ , and  $m_{ij}$  are elements in the weight, observed, and expected matrices, respectively.

### 3. METHODOLOGY

In this section we describe our algorithm framework to detect and diagnose the breast cancer metastases.

#### 3.1. Image Preprocessing

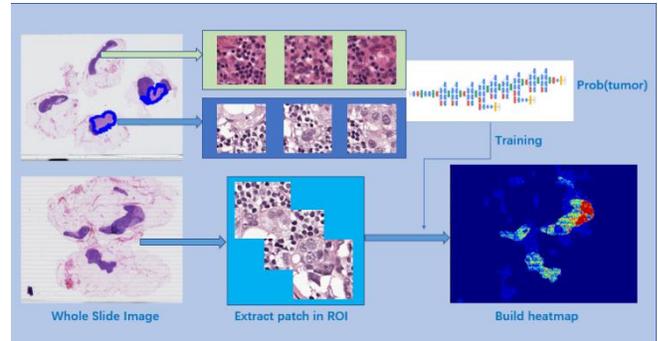


In general, the tissue region occupies partly the whole slide image. To pay attention to the region most likely to contain the cancer metastases., the first and significant step will be identifying the lymph node region. In addition to reducing the computing time and resources, the key is make our model to focus on the distinguishing cancer patches from lymph node tissue rather than other disturbing part(adipose tissue and so on).

To achieve this, we firstly get the low resolution image from WSI and the color space from RGB space to HSV space. We filter the hue, saturation, value, by specific range. Then we apply morphology dilation and erosion to the thresholding image. Finally we can find the contours and bounding box of our interested region.

This thresholding based segmentation method only find the lymph node region roughly. Later we use deep learning based segmentation(U-net, deeplab-v3, Mask Rcn) to get accurate lymph node region for training a fine-tuned convolution neural network.

#### 3.2. Automated Diagnosis Framework



Our Automated Diagnosis Framework consist of patch-based classification stage, slide-level and patient classification stage.

We extract randomly hundreds of thousands of small patches in the Training WSIs.' ROI found by segmentation. If the patch is located in tumor region annotated by the doctor,. We assign it the tumor label. Otherwise the normal label is assigned. The number of normal patch in the WSI is larger than the number of tumor patch. We control the number of normal patches extracted in each WSI and we extract all the tumor patch in the tumor region annotated by doctors to keep balance in positive and negative samples. By the way, we extracted uniformly in each area of WSI and each WSI of the whole training WSIs to cover the varied patches as far as possible .

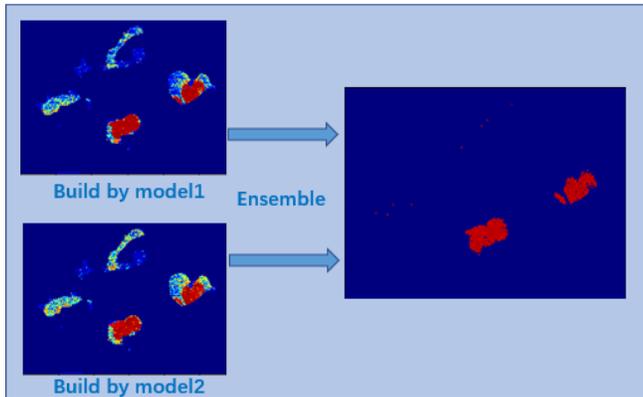
In recent years, deep CNNs have significantly improved accuracy on a wider range of computer vision tasks such as image recognition object detection, and semantic segmentation. We apply deep CNN to do patch-based classification task. We evaluated the predicting performance of different deep CNN architectures for the patch-based classification task. VGG16, Google Net, InceptionV3, ResNet101. The validation performance are shown below . ResNet101 and InceptionV3 achieve the best performance in these CNN architectures

<i>CNN Architecture</i>	<i>Validation Performance</i>
<i>VGG16</i>	93.51%
<i>Google Net</i>	97.12%
<i>InceptionV3</i>	97.80%
<i>ResNet101</i>	98.01%

Take the memory and computation resource into account, we finally choose the InceptionV3 as our final model.

We train a fine tuning model without imagenet weight running on 4 NVIDIA M40 GPUs. We use SGD with momentum 0.9 , and initial learning rate: 0.01, number of epoch to decay: 70 ,decay 0.001 and batch size 32 to train. After 18 hours, the loss converge to the optimal.

#### 3.3. Postprocessing



After we built the heatmap from the trained inceptionV3 model, the heatmaps generated doesn't look so good compared to the ground truth because of the false positive patch. To solve this problem, we collect false negative patch and false positive patch according to the difference between the heatmap and the ground truth.

The second model is trained by the same parameter. The only difference between model1 and model2 is that model2 is trained by the dataset enriched with hard negative patches and hard positive patches.

The final heatmap is built by the ensemble of two models.

```

if  $prob_{model1} > 0.9$  and  $prob_{model2} > 0.5$ 
    probability =  $prob_{model1}$ 
else
    probability = 0

```

### 3.4. Slide level and patient level classification

For slide level classification, we begin with the heatmap generated from original WSI. Due to the limitation of the number of WSI samples, we apply Artificial Feature Engineering on the heatmap to generate numeric features for training. We extract 33 numeric features such as the size of tumor region, the number of tumor regions, the long axis of tumor region and so on based on medical discrimination criteria. We need to be careful for overfitting because of many feature numbers and a small number of samples.

We apply Random Forest classifier to classify on WSI level. After getting the prediction on the WSI level, we can diagnose the patient pN-stage according to his 5 lymph node slides by the criteria mentioned in section 2.

## 4. EXPERIMENT RESULT

To achieve the best patient-level performance, we try different classifiers to do the final prediction. The ensemble prediction results are selected as our final prediction.

## 5. DISCUSSION

There is still a lot of work to finish and a huge space to improve. It is not the final version. To be continued.

## 11. REFERENCES

- [1] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. CoRR, abs/1409.1556, 2014. 3
- [2] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In CVPR 2015, 2015. 3
- [3] D. Wang, C. Otto, and A. K. Jain. Face Search at Scale: 80 Million Gallery. 2015. 3
- [4] D. L. Weaver, D. N. Krag, E. A. Manna, T. Ashikaga, S. P. Harlow, and K. D. Bauer. Comparison of pathologist-detected and automated computer-assisted image analysis detected sentinel lymph node micrometastases in breast cancer. Modern pathology, 16(11):1159–1163, 2003. 1
- [5] D. R. Cox. Regression models and life-tables. Journal of the Royal Statistical Society. Series B (Methodological), pages 187–220, 1972. 6
- [6] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In ICCV, pages 1026–1034, 2015. 2
- [7] P. J. Heagerty and Y. Zheng. Survival model predictive accuracy and roc curves. Biometrics, 61(1):92–105, 2005. 6
- [8] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz. Patch-based convolutional neural network for whole slide tissue image classification. In CVPR, pages 2424–2433, 2016. 1, 2
- [11] Camlyon17 <https://camelyon17.grand-challenge.org/>