DEEP LEARNING BASED PATIENT LEVEL BREAST CANCER METASTASIS CLASSIFICATION

North Star

NorthStar

ABSTRACT

Detection and classification of breast cancer metastasis in lymph nodes is a critical task towards improved patients' diagnosis and treatment. The diagnostic procedure for pathologists is tedious, time consuming and prone to errors especially due to small-sized metastases. Pathologists rely on TNM system in order to classify the extent of cancer spread. TNM takes into consideration the size of the tumour (T-stage), the spread to regional lymph nodes (N-stage) and whether it has metastasised to other body parts (M-stage). This short paper presents a method capable of performing the pN-stage (pathologic N-stage) classification of whole slide images (WSI), the major determinant of patient's prognosis and treatment. The pN-stage is predicted by combining convolutional neural network (CNN) based metastasis detector and a slide-level lymph node classifier module. The method was evaluated on Camelyon16 and Camelyon17 datasets which are challenging benchmark datasets.

Index Terms— Camelyon17, Convolutional neural network, Deep learning, Metastasis detection, Classification, pN-staging

1. INTRODUCTION

The breast cancer TNM (Tumor, Node, Metastasis) staging system[1] is the most common way that doctors stage breast cancer. The most critical part for the TNM stage classification process is to assess whether the breast cancer has spread to the regional lymph nodes (N-stage). However, due to time pressure, workload and other factors the diagnostic process followed by pathologists to detect metastases is prone to errors. This work aims towards reducing doctor's workload and help towards speeding up the diagnosis procedure.

Over the last few years, developments in hardware and more specifically improvements in the gpu-related technology, have allowed the extensive use of convolutional neural networks to address highly complex computer vision tasks[2]. Building on this notion, CNNs have been recently developed to perform computer assisted metastasis detection for breast cancer. Indicatively a relevant study [3], suggested the use of CNNs as a tool capable to detect not only breast cancer in sentinel lymph nodes but also other forms of cancer (such as prostate), while increasing the objectivity of diagnosis. Another study[4] suggested CNN based lymph node breast tumor detection framework which obtained state-of-the-art results on the Camelyon16 [5] dataset.

The pipeline presented here consists of three main modules:

1) a tissue region extraction module,

2) a CNN based metastasis detection module,

3) a WSI-level classification module.

More details regarding the aforementioned modules are provided in the following paragraphs.

The framework was evaluated on data from the contests Camelyon16 and Camelyon17[6], which included WSI from 5 different medical centers.

2. METHOD

In this study an efficient framework for pN-stage prediction based on patient's histological lymph node whole slide images is proposed,

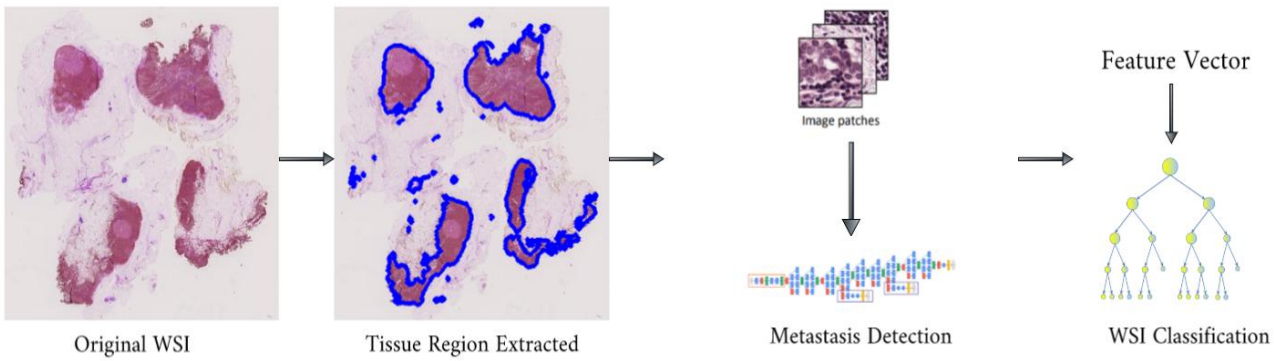

Figure 1: Algorithm Pipeline

Figure 1 illustrates an overview of the framework.

Initially the tissue region of each WSI is extracted and divided into smaller patches, which are fed as input to a convolutional neural network to produce a probability heatmap, indicative of the detected metastases. The heatmap is thresholded and a vector of features gets extracted in order to be utilized by the

following random forest classifier which predicts the WSI level label. Once every WSI of each patient is properly labeled, following the set of rules described in Camelyon 17 each patient gets classified to the respective pN-stage.

An important aspect of the method is that during training a scheme of hard negative mining is utilized involving humans in the loop. More specifically, after careful visual inspection of the patches it was chosen whether they would be added to the existing set of patches, for the following retraining step or not. This way patches with possibly missanotated ground truth were excluded, allowing the network to reach higher accuracy levels.

2.1. Patch Extraction

An average WSI is approximately 200000 x 100000 pixels on the highest resolution level (zero level) and approximately 7.5 GB. Hence processing the whole image at once, is highly

inefficient due to the enormous computation time required. The first step of the process is to

isolate the part of the WSI that contains potentially useful information, the tissue region.

This way the vast percentage of the image which is assumed to be background is not used either for training purposes or for the inference process. Recent studies[7],[8] indicated that the tissue regions can be efficiently extracted using a variety of thresholds ranging such as Otsu thresholding[9] of the respective grayscale image or thresholding of the image in other color-spaces (i.e. HSV, HSL etc) which can potentially reveal some additional useful pieces of information.

A worth mentioning observation made by visual inspection is that metastatic regions can be located even at the edge of the tissue regions. Therefore, the need for a highly accurate tissue region extraction approach is clear. After extensive experimentation on WSIs from both Camelyon16 and Camelyon17 it was found that the use of Otsu thresholding

on the grayscale image followed by a series of morphological image operations allows the extraction of a quite detailed tissue mask from the original WSI. For the next step of the pipeline, the tissue mask extracted gets divided into patches of 299*299 size.

2.2. Data Augmentation

Having acquired images from 5 different centers the classification of the WSIs has been proved to be a problem that requires a robust artificial intelligence approach capable of generalizing well. The variety of hematoxylin and eosin (H&E) stained color due to the different chemical preparation process performed by each center, was addressed with an extensive color augmentation scheme. More specifically random hue, saturation, brightness, and contrast changes were introduced to each patch included in the training process of the CNN. In addition, taking advantage of the fact that the histopathology images exhibit rotational symmetry, the data were augmented by random rotations and flips.

2.3. Convolutional Neural Network (CNN)

For the actual training Tensorflow and an architecture based on Inception-V1 was used. The network converged after 10 epochs.

2.4. Heatmap Generation

The CNN was used to produce 299*299 pixels probability heatmaps, which were combined to produce a full WSI dimensions heatmap at level 7. These heatmaps were thresholded with t=0.85, in order to extract a list of more morphological features per WSI, which were fed as input to a random forest classifier that assigned them a "negative", "itc", "micro" or "macro" label. The list of features chosen includes but is not limited to the following:

Features extracted from thresholded probability heatmap		
area of ellipse fitting the biggest tumor	major axis of ellipse fitting the biggest tumor	average probability of biggest tumor
maximum probability of the WSI	mean of the probabilities in the WSI	variance of the probabilities in the WSI

2.5. Whole Slide Image Classification and pN-staging

The feature vector extracted from the previous step is fed as input to a random forest classifier in order to assign a WSI-level label. This extra module was used in order to compensate for the imperfections of the extracted heatmap and the errors that are introduced to the measurements of the major axes of the segments included. The final step is to use the set of rules provided by Camelyon17 in order to extract the patient level pN-stage.

3. RESULTS

Mean Intersection Over Union (IoU) in patch level is 0.71 from validation slides for heatmap threshold equal to 0.85. Smaller threshold values were also tested but they allowed more false positives detections.

4. DISCUSSION

The prevalent approaches of pN-stage predictions focus on patch-level classifications, due to the size of the whole slide images which does not allow to perform the training at the original size, because of hardware restrictions. At the same time the number of the fully annotated whole slide images is small making making the training process harder and the danger of overfitting imminent. For the aforementioned reasons, techniques like l2 normalization, dropout and batch normalization were integral parts of the proposed method.

A critical limitation of this approach is that, despite the fact that the classifier reaches very high accuracy scores, false positive alarms still occur. Hence, WSIs which contain only ITCs are rarely labeled properly as a number of ITCs are often miss-labeled at segment level as "micros". Thus the focus of this study was to achieve the highest possible accuracy for negative and micro, macro metastases.

5. CONCLUSION

A deep learning framework was proposed in this study in order to predict the pN-stage from whole slide histopathology images, performing CNN based metastasis detection and random forest based lymph node classification.

The performance of the method is proven to be competitive with other state of the art algorithms, after being tested on the Camelyon dataset.

Future work will include creating a new end-to-end deep learning framework possibly fully automated aiming to achieve better results for the pN-stage prediction from WSI.

6. 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] Razzak, Muhammad Imran, Saeeda Naz, and Ahmad Zaib. "Deep learning for medical image processing: Overview, challenges and the future." Classification in BioApps. Springer, Cham, 2018. 323-350.

[3] Litjens, Geert, et al. "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis." Scientific reports 6 (2016): 26286.

[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] "Camelyon 2016", https://camelyon16. grand-challenge.org/, Accessed: 02/07/2019.

[6] "Camelyon 2017", https://camelyon17.grand-challenge.org/, Accessed: 02/07/2019.

[7] Bejnordi, Babak Ehteshami, et al. "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer." Jama 318.22 (2017): 2199-2210.

[8] Bandi, Peter, et al. "From detection of individual metastases to classification of lymph node status at the patient level: the CAMELYON17 challenge." IEEE transactions on medical imaging 38.2 (2018): 550-560.

[9]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.