Tumor Detection in Breast Histopathology Images via modified Faster-RCNN

○Talha Ilyas[1], 김상철[2*], 김형석[3*]
1)○ 전북대학교, 전자 정보공학부 (TEL: 010-9761-2736; E-mail: talha@jbnu.ac.kr)
2)* 전북대학교 부설 지능형로봇연구소 (TEL: 063-219-5282; E-mail: sckim7777@jbnu.ac.kr)
3)* 전북대학교, 전자 정보공학부, 전북대학교 부설 지능형로봇연구소
|(TEL: 063-270-2477; E-mail: hskim@jbnu.ac.kr)

Abstract Biopsied tissue detection and classification within Breast Histopathology Images is a fundamental prerequisite to estimate the aggressiveness of breast cancer. The development and fully automated pipelines for tissue detections and classification enables the analysis of thousands of tissues within a whole slide histology image, which opens possibilities for analysis and prognosis of breast tumor. There are multiple annotated histology datasets available for evaluating the performance of machine learning models. The number of samples in these datasets is quite limited and usually the annotations provided are in the form of pair of points which points to the center of different types of cells. Most of the works in this field approach this problem by cropping a patch of the WSI usually 50x50 pixels (centered at given point), and then classify these patches with a simple classifier CNN. In this work we propose a method of converting the provided annotations (center points) into bounding box annotations. Then we use Faster-RCNN to detect and classify different types of cells in the WSI in a fully automated pipeline. We also propose data augmentation technique to increase the dataset size for Breast Histology images. Our proposed approach showed an average precision of 70.34% for classification and detection of tumor tissues.

Keywords Object Detection, Breast Histopathology Images, Cancer Detection, Tumor cells, Deep Learning.

1. Introduction

ISTOPATHOLOGICAL tissue analysis by a pathologist determines the diagnosis and prognosis of most tumors, such as breast cancer. To estimate the aggressiveness of cancer, a pathologist evaluates the microscopic appearance of a biopsied tissue sample based on morphological features which have been correlated with patient outcome. The task associated with this dataset is to automatically classify histological structures in these hematoxylin and eosin (H&E) stained images into six classes: (a)mitosis, (b) apoptosis, (c)tumor, (d) non-tumor, (e)lumen, (f) non-lumen. Manually spotting and annotating the affected area(s) on histopathology images with high accuracy is regarded as the gold standard in cancer diagnosis and grading, but it is also a time-consuming and tedious task that requires considerable effort, expertise, and experience of pathologists. These skills are mostly gained over time by analyzing more cases. Whereas this visual interpretation has strict guidelines, it brings a certain subjectivity to the histological analysis, and therefore leads to inter/intra-observer variability. Besides, the variability in size, shape, location, texture of nuclei, turn automated detection into a tedious and more difficult task.

In this paper we use the BreCaHAD dataset [1] to perform experiments and analysis the performance of our proposed approach. The annotations provided with the BreCaHAD are json files for each image the corresponding file contains the normalized center points of the different cell types.

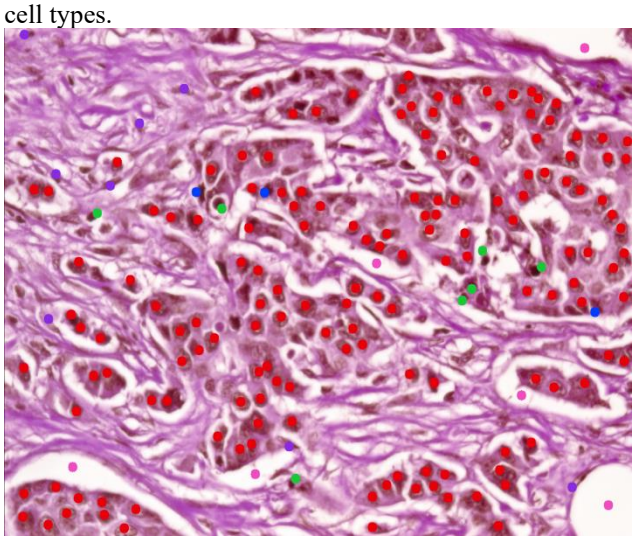

Figure 1: Sample image of BreCaHAD dataset with corresponding annotations, different colored dots show different cell types.

In this work we converted those center points into the un-normalized bounding boxes with width and height of (45,45) pixels, we choose this area because with this size the box completely contains a single cell [2]. Usually, this type of problem is solved via FCNs to do binary segmentation and then classify the segments or just by cropping a patch of the

WSI (whole slide image) and the classify each patch separately [2, 3]. So, we propose a method to solve this problem via single detection pipeline (SSD, Faster_RCNN etc.). But for that first we need to convert this dataset into a proper detection dataset format, as the dataset is a main component in training a CNN. So, we will modify the provided annotations and convert them in to bounding boxes annotation.

2. Generating Dataset

The dataset consists of 162 breast cancer high resolution histopathology images. The dataset includes various malignant cases. Firstly, we will pre-process and simplify the data such that our detection network can extract more meaningful features. Image 1(a) shows a sample image of the BreCaHAD data along with the provided annotations. The cell centers are provided so I put a circle at each cell's center for easy visibility. Different colors show different cell types.

2.2. Data Preprocessing

Firstly, we converted the given annotations into the bounding boxes of height and width of 45, 45 as this area completely encapsulates a cell [3, 4]. I saved the annotation in .xml format for later use in data pre-processing. A sample image is shown Figure 2.

As shown in Figure 2 we can clearly see that the images are densely annotated, and the resolution is very high. Moreover, a lot of cells of different type are cluttered together, which will make it difficult for the network to precisely localize and classify them. As the cells already have very little inter-class variations, we need to find a way to process the data which will increase spatial resolution of the data. But if we directly upscaled an image of resolution 1360x1024 to an even bigger resolution the computation time of CNN will explode. So, one solution could be resizing all the images. But resizing them will reduce the WSI resolution, we might lose a lot of useful information. So, to cope with this problem, we can do upscale cropping or also called SSD crop data augmentation.

※ 이 논문은 2021 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No.NRF-2019R1A6A1A09031717). 또한, 본 논문은 농촌진흥청 공동연구사업(과제번호 : PJ015720)의 지원에 의해 이루어진 것임

2.1. Data Augmentation (SSD Crop)

Instead of resizing the images directly first we will crop a part of the image (with predefined aspect ratio) and then rescale it to the input size [5]. This might sound simple but it's not, because you don't only have to resize the image but also you will have to update its corresponding annotations in the .xml file to the new rescaled coordinates. So, from a single image we will get 9 crops form different parts (i.e., 'center', 'left-top', 'left-center', 'left-bottom', 'center-top', 'center-bottom', 'right-top', 'right-center', 'right-bottom') as shown in figure below.

Shaded regions in Figure 2 show a single crop. Each crop has
the same aspect ratio as the original image and is have 10~15% overlap with its adjacent neighbor. Using this technique, we can avoid reduce in WSI resolution and quality. Using SSD crop will give us two major benefits;

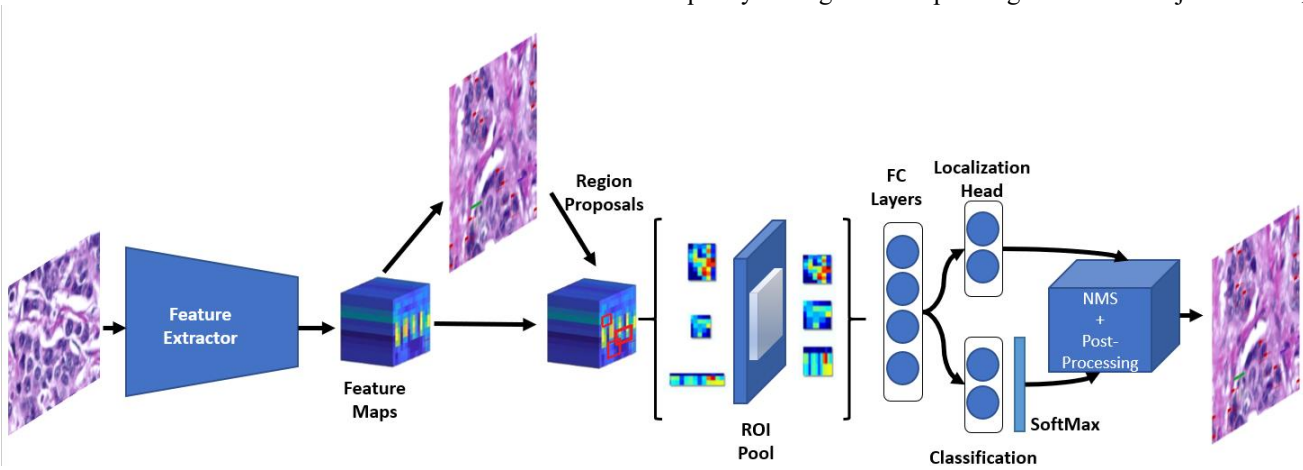

Figure 2: Proposed Network Architecture. We use [ResNet, Inception-Net and Xception-71] as feature extractor.

(a)spatially enhance the data for the CNN for better performance (b) increase the data size from 162 original images to 1458 SSD crop images. Note: This
data is not the rotated or flipped version of the image (as in typical data augmentation), but each image in data is unique because we cropped the original image from different parts and then rescaled it to make the new set.

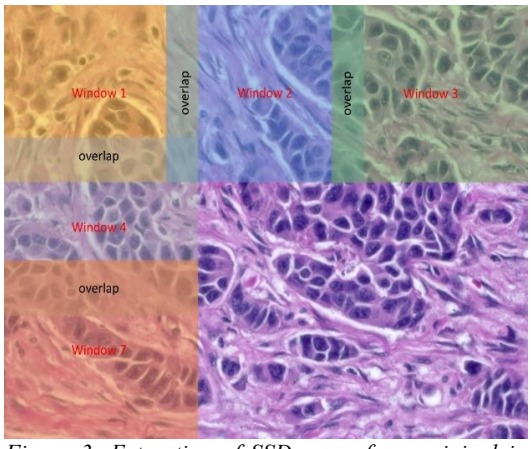

Figure 3: Extraction of SSD crops form original image to increase the dataset size and reduce the computational complexity. Shaded regions show a single crop. Each crop has the same aspect ratio as the original image and is have 10~15% overlap with its adjacent.

Figure 3 shows SSD crops of original ones along with their scaled annotations. Form the following images we can clearly see that we have simplified the data quite a bit.

3. SYSTEM OVERVIEW AND DETECTION ARCHITECTURE

Our work aims to identify six classes of tissues in breast histopathology images using Deep Learning as the main body of the system. A general overview of the system is presented in Figure 4. Following we describe in detail each component of the proposed approach.

3.2. Detection Architecture

We decided to use the Faster-RCNN [6] meta-architecture because of its superior ability to detect region of interests in complex and cluttered scenarios. As for the network back bone, seeing the complexity of dataset, we decided to combine two state of the art Feature extractors i.e., ResNet [7] and Inception-Net [8] to get best performance. We used 150-layer Inception-ResNet as the backbone feature extractor. We perform the non-maximum suppression (NMS) and other post-processing on networks output to get visual results. The full network architecture is shown in Figure 2.

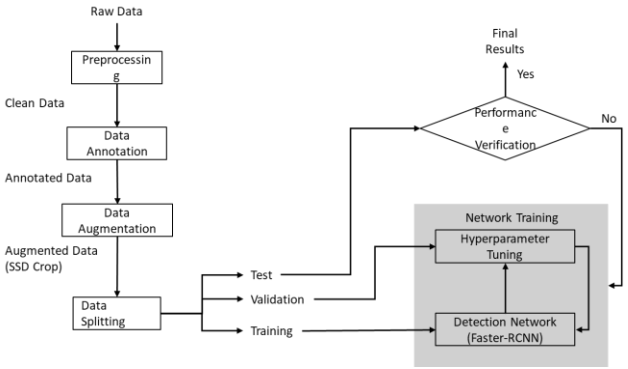

Figure 4: System overview of the proposed deep-learning-based approach for breast tumor detection.

4. Quantitative Results

We evaluate our model on both COCO [6] and PASCAL-VOC [9] style mAP metric. The Average Precision is the area under the Precision-Recall curve for the detection task. As in the Pascal VOC Challenge, the AP is computed by averaging the precision over a set of spaced recall levels [0, 0.1,…, 1], and the mAP is the AP computed over all classes in our task.

$$AP = \frac{1}{11}\sum_{r\in\{0,0.1,\ldots,1\}} p^r_{interp} \quad (1)$$

$$p_{interp}(r) = max_p(r) \quad (2)$$

where p(r˜) is the measure precision at recall r˜. Next, we compute the mAP averaged for an IoU = 0.5 (due to the complexity of the scenarios).

5. Results and Discussion

For our experiments we used a batch size of 8, and learning rate used was Cosine Learning Decay, with warmup steps of 3000 and trained the network for a total of 100000 iterations. The dataset is highly imbalance in favor of tumor class Figure 5. So, to make the learning easy we made the following modifications. We removed the lumen and non-lumen class because these classes are not useful in downstream analysis of breast cancer. As for non-mitosis class there is no annotations available in the data as shown in bar graph below.

5.2. Quantitative Results

Results of our experiments are reported in Table 1. As can be seen from table that using our proposed approach we were successfully able to detect the tumor cells form breast histopathology images. The results also give us meaningful insight to the problem, that when we use a stronger backbone (CNN with high classification accuracy) then the mAP is also high. Which means that network is having more difficulty classifying the detecting tissues rather then having difficulty in detecting them. Which means that our algorithm can greatly benefit form a stronger backbone like FPN.

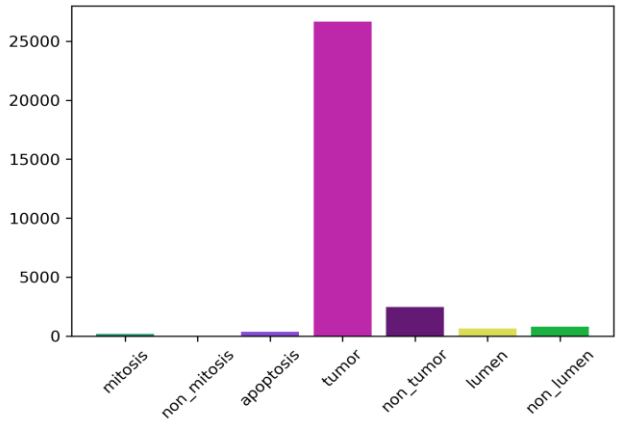

Figure 5: The bar graph shows the number of instances of each class present in the dataset.

5.3. Visual Results

Figure 6 shows visual results of our algorithm. Form visual results a lot of interesting things come into lig ht. First is that network was able to classify and loca lize different types of cells very well with high preci sion and confidence. Second, if you look at the third-row predictions, we can see that the network has det ected a lot of tumor cells in this cropped image patc h even though none of them is present in the ground truth image. One potential explanation for this may be because, as we have shown earlier, the WSI is ve ry high resolution and can have hundreds of instances clustered together in a small area of WSI. Annotatin g such high-density images with high precision is ver y exhausting and requires a lot of time and human la bor. Moreover, because of the difficulty of the task, t here is a high likelihood of human error, just as seen in row 3 of the visuals.

TABLE 1. Quantitative results of proposed approach.

Architecture	Backbone	$AP^{Apoptosis}$	AP^{Tumor}	$AP^{Mitosis}$	$AP^{Non-Tumor}$	mAP
Faster-RCNN	ResNet-50	33.45	60.58	29.84	22.40	36.56
Faster-RCNN	Inception-ResNet-50	38.6	66.18	37.94	24.60	41.83
Faster-RCNN	Xception-71	41.3	70.39	40.97	30.5	45.79

6. Conclusion

As can be seen from the visual results that network was ab le to detect the tumor cells with high precision, but in this case one would usually want high recall rather than high p recision. So, that none of the tumor cell is ignored by the n etwork, this can be an interesting direction to work on in f uture. Moreover, our algorithm can also benefit form a str onger backbone architecture.

7. References

1. Aksac, A., et al., *BreCaHAD: a dataset for breast cancer histopathological annotation and diagnosis.* BMC research notes, 2019. **12**(1): p. 1-3.
2. Roy, K., et al., *Patch-based system for Classification of Breast Histology images using deep learning.* Computerized Medical Imaging and Graphics, 2019. **71**: p. 90-103.
3. Wang, Z., et al. *Classification of breast cancer histopathological images using convolutional neural networks with hierarchical loss and global pooling*. in *International Conference Image Analysis and Recognition*. 2018. Springer.
4. Xie, J., et al., *Deep learning based analysis of histopathological images of breast cancer.* Frontiers in genetics, 2019. **10**: p. 80.
5. Liu, W., et al. *Ssd: Single shot multibox detector*. in *European conference on computer vision*. 2016. Springer.
6. Ren, S., et al., *Faster r-cnn: Towards real-time object detection with region proposal networks.* IEEE transactions on pattern analysis and machine intelligence, 2016. **39**(6): p. 1137-1149.
7. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
8. Szegedy, C., et al. *Inception-v4, inception-resnet and the impact of residual connections on learning*. in *Proceedings of the AAAI Conference on Artificial Intelligence*. 2017.
9. Everingham, M., et al., *The pascal visual object classes (voc) challenge.* International journal of computer vision, 2010. **88**(2): p. 303-338.

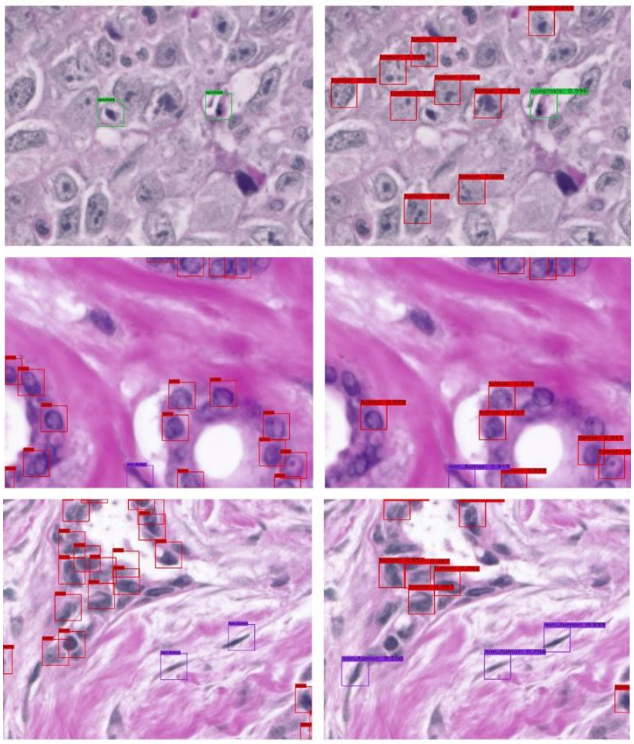

Figure 6: Visual results of proposed approach. Left column shows ground truths and Right columns shows predictions.