

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/340998089>

[Rank 6] Mask-RCNN for Cell Instance Segmentation

Article in IEEE Transactions on Medical Imaging · April 2020

CITATIONS

0

READS

893

5 authors, including:



Sihang Zhou

National University of Defense Technology

41 PUBLICATIONS 426 CITATIONS

[SEE PROFILE](#)



Qian Wang

Nanjing Tech University

193 PUBLICATIONS 4,994 CITATIONS

[SEE PROFILE](#)



Neeraj Kumar

University of Alberta

70 PUBLICATIONS 728 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Neuroimaging Processing, Analysis and its application in Dementia Identification [View project](#)



polyoxometalate [View project](#)

Mask-RCNN for Cell Instance Segmentation

Xuhua Ren¹#, Sihang Zhou²#, Dinggang Shen² and Qian Wang¹

¹ Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China

² Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599
renxuhua@sjtu.edu.cn

Abstract. We proposed an automatic nucleus segmentation algorithm of H&E stained tissue microscopy images. Mask-RCNN is a recently proposed state-of-the-art algorithm for object detection and object instance segmentation of natural images. In this paper, we demonstrate that Mask-RCNN can be used to perform highly effective and efficient automatic segmentation of H&E microscopy images for cell nuclei. We propose a novel MASK Non-maximum suppression (NMS) module which can automatically ensemble classifiers results and increase the robustness of model.

Keywords: Cell Segmentation, H&E stained, Mask-RCNN, Deep Learning

1 Introduction

Cell instance segmentation is an important task in medical image analysis involving cell level pathology analysis. In H&E stained microscopy images, this task is challenging for various reasons, for example the relatively large variation in the intensity of captured signal, and the ambiguity boundary information when separating neighboring cells. It requires careful model configurations to ensure its robustness to capture certain feature information of cells in the images such as intensity, shape, and size. Manual segmentation can be time consuming. In contrast, deep learning-based approaches have shown the great power of automatic extraction and selection of cell image features. In this paper, we present a novel Mask-RCNN [1] algorithm to solve the problem of cell segmentation in H&E stained microscopy images.

The Mask-RCNN model was developed in 2017, which is extended from the Faster-RCNN [2] model for semantic segmentation and object instance segmentation of images. Mask-RCNN has shown its effectiveness among all existing single-model entries on every task in the 2016 COCO Challenge. Mask-RCNN relies on a region proposal which are generated via a region proposal network. Mask-RCNN follows the Faster-RCNN model of a feature extractor followed by this region proposal network, then followed by an operation known as ROI-Pooling to produce standard-sized outputs suitable for input to a classifier, with two important modifications. First, Mask-RCNN replaces the imprecise ROI-Pooling operation with an operation called ROI-Align that allows very accurate instance segmentation masks to be constructed; Sec-

ond, Mask-RCNN adds a network head (a small fully convolutional neural network) to produce the instance segmentations; c.f. Figure 1.

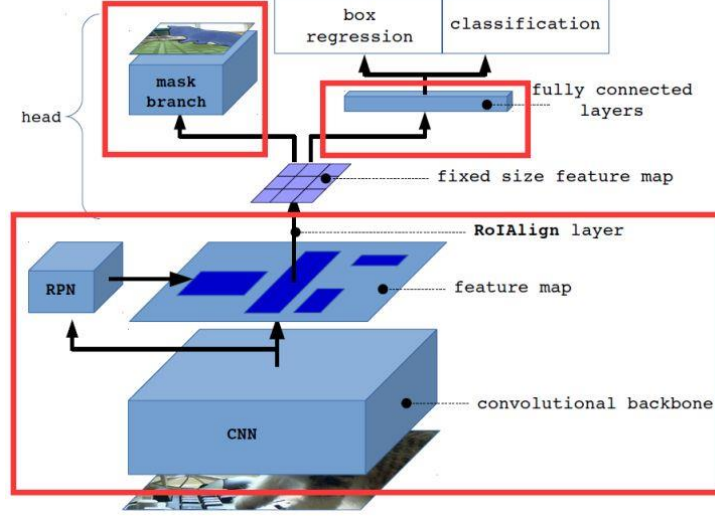


Fig. 1. The Mask-RCNN model.

2 Materia and Method

2.1 Dataset

The dataset for this challenge was obtained by carefully annotating tissue images of several patients with tumors of different organs and were diagnosed at different hospitals. The H&E stained tissue images were captured at 40x magnification from TCGA archive. H&E staining is a routine protocol to enhance the contrast of a tissue section and is commonly used for tumor assessment (grading, staging, etc.).

2.2 The Proposed Network

In this paper, we use a Mask-RCNN model with a ResNet-101[3] pyramid network backbone. It is based on an existing implementation by Matterport¹ and which is itself based on the open-source libraries Keras [4] and Tensorflow [5]. Rather than training the network end-to-end from the start, we initialize the model using weights obtained from pre-training on the MSCOCO dataset and proceed to train the layers in two stages: First, the model only trains the network heads, which are randomly initialized, then training the upper layers of the network, and then reducing the learning rate by a factor of 10 and training end to end. In total we trained 300 epochs using sto-

¹ https://github.com/matterport/Mask_RCNN

chastic gradient descent with momentum of 0.9. We use a batch size of 5 on a single NVIDIA Titan XP GPU.

Besides, we have implemented the following pre-processing works to ensure the validity of the experiments: We fill holes in masks by image morphology. We also split nuclei masks that are fused by applying morphological erosion and dilation. Zero mean unit variance normalization also utilized in each image. To help avoid overfitting, we implemented data augmentation using random crops, random rotations, gaussian blurring, random horizontal and vertical flips.

There are several changes made in the Mask-RCNN model to improve the segmentation performance: We reduced RPN (region proposal network) anchor sizes, since the nuclei are mostly small; Increased number of anchors to be used, since the nuclei are small and can be found anywhere on an image; Increased maximum number of predicted objects, since an image can contain 1000 or more nuclei. Increased POST_NMS_ROIS_TRAINING to get more region proposals during training. Added extra parameter DETECTION_MASK_THRESHOLD to model configuration. Cropped images and masks to 600x600.

Furthermore, to improve the robustness on segmenting cells with irregular shape (like, strip) and size (smaller than 6 pixels in radius), we introduce a multi-task U-Net like network². The code of the network is based on the release version on github. Specifically, the backbone of the u-net-like network is a pre-trained ResNet 101. Three tasks are conducted in the network. In the first task, we do the common segmentation, and try to segment all the fore-ground regions. In the second task, we segment the adjacent boundaries of between cells. In the third task, we conduct erosion operation to the instance masks and construct the interior portions of each cell, then do the segmentation on these interiors. The three tasks share the same back bone but different decoders. In the testing time, the foreground segmentation, adjacent boundaries and the interior of cells are inputted to a watershed algorithm for final segmentation.

For post-processing, we combined predictions from five-fold cross training models: took unions of masks with maximum overlap and removed false positive masks with small overlap. We called this module as MASK-NMS. Moreover, our NMS starts with a list of segmentation results I with scores S . After selecting the segmentation with the maximum score M , it removes it from the set I and appends it to the final segmentation set D . It also removes any segmentations which has an overlap greater than the threshold N_i compared with M in the set I , we address intersection over union (IOU) as overlap metric. This process would repeat until set I become empty. This module could ensemble multi-models results together and reduce false positive or false negative situation.

We also tried some unsuccessful strategies. E.g., Trained the model with Dice Coefficient Loss instead of default binary cross-entropy loss for the masks heads. Trained with random Gaussian noise for image augmentation. It hurt overall model performance. Tried assembling actual image predictions with horizontal and vertical flip predictions. Used non-maximum suppression for removing overlaps. Did not

² <https://github.com/samuelschen/DSB2018>

improve prediction accuracy on the validation set. Trained end-to-end without initializing with pre-trained ImageNet weights. It would be worse than pretrained model. Trained on preprocessed images with adaptive histogram equalization (CLAHE). The model performed way worse. Trained with mask dilations and erosions, this did not have any improvement in the segmentation in my experiments. And soft-NMS did not improve accuracy also.

There are some strategies that we didn't have time to try. E.g., Hyperparameter search on thresholds and network architecture. Different layer normalization techniques, with batch size more than one image at a time. Augmentation smoothing during training, i.e., Increase the noise and augmentation slowly during the training phase.

3 Experimental Results

The proposed Mask-RCNN model with ResNet-101 backbone obtains an average mask intersection over union (IOU) of 45.02% on the five-fold validation dataset. The image with nuclei detections and segmentations are illustrated in Figure 2.

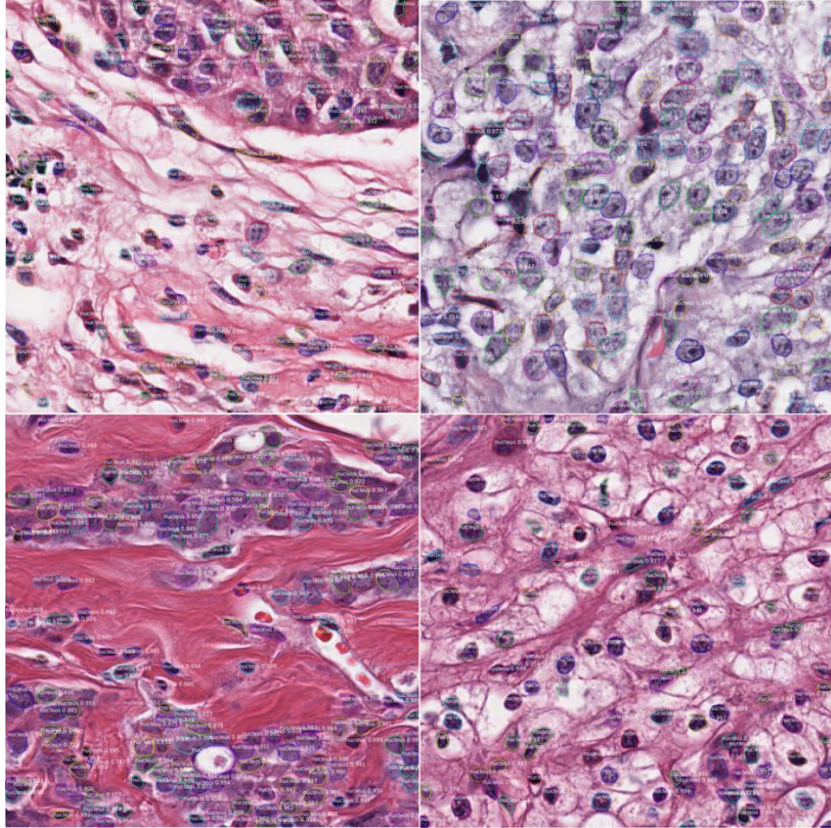


Fig. 2. Segmentation result in validation dataset.

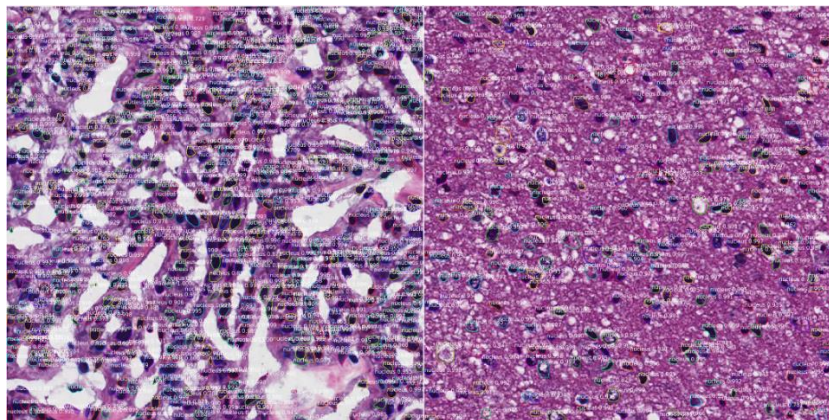


Fig. 3. Segmentation result in test dataset.

Conclusion

Cell segmentation is an important step for cell-level analysis of biomedical images. In this paper we demonstrate that the Mask-RCNN model, can be used to produce high quality results for the challenging task of segmentation of nuclei. We also designed a MASK-NMS for model ensemble which could reduce false positive or false negative situation.

References

1. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn. In: Computer Vision (ICCV), 2017 IEEE International Conference on, pp. 2980-2988. IEEE, (Year)
2. Girshick, R.: Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp. 1440-1448. (Year)
3. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. (Year)
4. Chollet, F.: Keras. (2015)
5. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M.: Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 (2016)