Colon Nuclei Instance Segmentation
using a Probabilistic Two-Stage Detector

Pedro Costa[1, 2], Yongpan Fu[3], João D. Nunes[1,2], Aurélio Campilho[1,2], and Jaime S. Cardoso[1,2]

[1]*Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal*
[2]*Institute for Systems and Computer Engineering Technology and Science (INESC TEC) 4200-465 Porto, Portugal*
[3]*Czech Technical University, Faculty of Electrical Engineering, Prague, Czesh Republic*

Abstract

Cancer is one of the leading causes of death in the developed world. Cancer diagnosis is performed through the microscopic analysis of a sample of suspicious tissue. This process is time consuming and error prone, but Deep Learning models could be helpful for pathologists during cancer diagnosis. We propose to change the CenterNet2 object detection model to also perform instance segmentation, which we call SegCenterNet2. We train SegCenterNet2 in the CoNIC challenge dataset and show that it performs better than Mask R-CNN in the competition metrics.

1. Introduction

Current cancer diagnosis is very time consuming, having very low throughput, and is prone to high inter and intra observer variability [3]. Parhologists start by performing a biopsy on the suspicious tissue, then the tissue is stained using Hematoxylin and Eosin, and finally the tissue is analyzed under a microscope.

Pathologists look for diverse histological properties of WSIs while searching for signs of cancer. These properties include the organization of cell nuclei, their density and other nuclei morphological features [2]. These features can then be used to predict survival [1, 7] or characterize the disease [7].

Deep Learning models could be used for the detection of individual nuclei and the classification of the type of cell they belong to. However the development of a Deep Learning model that detects and segments cell nuclei in WSIs is challenging. In particular, we must consider memory constrainst imposed by the high dimensional WSIs. Besides, each image contains a high density of cell nuclei, tipically more than 100, which is by far superior than the typical number of objects that standard object detection models are trained on.

In this work we implement an instance segmentation model, that simultaneously detects and segments cell nuclei in crops of WSIs. We propose to use CenterNet2 [10], a probabilistic two-stage object detection model. This model allows the reduction of proposals from the Region Proposal Network (RPN), which could be important in this application where each image has many objects. We also updated the original model to perform segmentation. Additionally, we propose a novel Region of Interest (RoI) Head that explores the spatial relationships between objects in the image.

2. Method

2.1. CenterNet2

The main contribution CenterNet2 was a probabilistic interpretation of two-stage object detection models. This novel interpretation allowed the proposal of a new training objecive that ties the two-stages together.

The goal of an object detection model is to produce a set of K bounding box detections, with an associated class distribution $P(C_k = c)$ for classes c including the background class. Two-stage detectors approach this problem by first modeling a class-agnostic object likelihood $P(O_k)$ in the first stage and a conditional categorical classifcation $P(C_k|O_k)$ in the second stage. The joint class distribution is as follows:

$$P(C_k) = \sum_o P(C_k|O_k = o)P(O_k = o), \qquad (1)$$

where o is 0 when the object belongs to the background and 1 when it is a positive detection.

For positive examples, object detection models can be trained by maximizing the following objective:

$$logP(C_k) = logP(C_k|O_k = 1) + logP(O_k = 1). \quad (2)$$

However, for negative examples, the objective does not factorize, since we may have positive detections in the first

stage that are classified as background in the second stage:

$$logP(bg) = log(P(bg|O_k = 1)P(O_k = 1) + P(O_k = 0)). \quad (3)$$

The authors proposed to maximize two lower bounds of $logP(bg)$:

$$logP(bg) \geq P(O_k = 1)log(P(bg|O_k = 1)), \quad (4)$$

$$logP(bg) \geq log(P(O_k = 0)). \quad (5)$$

They jointly optimize these two lower bounds and show improved performance.

The first stage of the detector uses CenterNet [11], but CenterNet2 uses a ResNet-FPN [6] backbone. Then, for the second stage, we use a RoIAlign followed by a standard RoIHead to regress the bounding box location and object class.

2.2. Instance Segmentation

CenterNet2 original paper only supports object detection, therefore, we updated CenterNet2's RoI Head to also include mask prediction, following Mask R-CNN's [5] approach.

For that, we added 3 additional Fully-Connected layers that map the region of interest feature vector into a fixed size mask of 14×14px. We only predict a single segmentation mask, independent of the object class. The ground-truth mask is resized to the resolution of the predicted mask before applying an IoU loss. We call this modified CenterNet2 as SegCenterNet2.

2.3. Implementation Details

We used Detectron2 [9], a PyTorch [8] framework to implement and train our models. We trained the models for 15000 iterations with a batch size of 8. We used a warmup learning rate scheduler, where the learning rate is linearly increased during 2000 interations until reaching its final value of 0.02. The learning rate was divided by 10 after 12500 iterations and then again after 14000. The Adam optimizer was used with gradient clipping.

The input images were resized to 800×800px to increase the spatial resolution of the FPN. We used random horizontal flips as data augmentation. At test time, detections with a score higher than 0.5 were considered.

3. Evaluation

3.1. Dataset

In this work, we used the Colon Nuclei Identification and Counting Challenge (CoNIC) [4] dataset. The dataset consists of 4981 non-overlapping Hematoxylin and Eosin stained histology patches obtained in 5 different centers.

Method	PQ	Multi-PQ	DQ	SQ
HoverNet	**0.6149**	**0.4998**	-	-
SegCenterNet2	0.6067	0.4880	0.7846	0.7400

Table 1. Comparison between our SegCenterNet2 model with HoverNet.

Each cell nucleus belongs to one of 6 classes: neutrophil, epithelial, lymphocyte, plasma, eosinophil, and connective. We used the same split that the authors of the competition provided, making sure that images from the same patient are all in the same set.

3.2. Instance Segmentation

To evaluate our model, we use the competition metrics: PQ and multi-PQ. In Table 1 we can see that SegCenterNet2 performs slighlty worse than the baseline, however, by only a small margin.

References

[1] Najah Alsubaie, Korsuk Sirinukunwattana, Shan E Ahmed Raza, David Snead, and Nasir Rajpoot. A bottom-up approach for tumour differentiation in whole slide images of lung adenocarcinoma. In *Medical Imaging 2018: Digital Pathology*, volume 10581, page 105810E. International Society for Optics and Photonics, 2018. 1

[2] Teresa Araújo, Guilherme Aresta, Eduardo Castro, José Rouco, Paulo Aguiar, Catarina Eloy, António Polónia, and Aurélio Campilho. Classification of breast cancer histology images using convolutional neural networks. *PloS one*, 12(6):e0177544, 2017. 1

[3] Joann G Elmore, Gary M Longton, Patricia A Carney, Berta M Geller, Tracy Onega, Anna NA Tosteson, Heidi D Nelson, Margaret S Pepe, Kimberly H Allison, Stuart J Schnitt, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *Jama*, 313(11):1122–1132, 2015. 1

[4] Simon Graham, Mostafa Jahanifar, Quoc Dang Vu, Giorgos Hadjigeorghiou, Thomas Leech, David Snead, Shan E Ahmed Raza, Fayyaz Minhas, and Nasir Rajpoot. Conic: Colon nuclei identification and counting challenge 2022. *arXiv preprint arXiv:2111.14485*, 2021. 2

[5] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017. 2

[6] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017. 2

[7] Cheng Lu, David Romo-Bucheli, Xiangxue Wang, Andrew Janowczyk, Shridar Ganesan, Hannah Gilmore, David Rimm, and Anant Madabhushi. Nuclear shape and orientation features from h&e images predict survival in early-stage estrogen receptor-positive breast cancers. *Laboratory investigation*, 98(11):1438–1448, 2018. 1

[8] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. 2

[9] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. https://github.com/facebookresearch/detectron2, 2019. 2

[10] Xingyi Zhou, Vladlen Koltun, and Philipp Krähenbühl. Probabilistic two-stage detection. *arXiv preprint arXiv:2103.07461*, 2021. 1

[11] Xingyi Zhou, Dequan Wang, and Philipp Krähenbühl. Objects as points. *arXiv preprint arXiv:1904.07850*, 2019. 2