Multi-Class Cell Detection Using Spatial Context Representation – Supplemental Material –

Shahira Abousamra, David Belinsky, John Van Arnam, Felicia Allard, Eric Yee,
Rajarsi Gupta, Tahsin Kurc, Dimitris Samaras, Joel Saltz, Chao Chen
Stony Brook University
Stony Brook, NY 11794, USA

A. Model Architecture

Fig. 1 shows details of the proposed deep neural network architecture.

B. Datasets

We provide more statistics of the datasets in Table 1. In Figure 2, we provide sample patches and annotations from the three datasets. It should be noted that even point annotation is expensive. It takes around 30 minutes to annotate a patch by our pathologist annotator.

C. Baselines

We provide details of the baseline algorithms.

HoverNet-Weakly. Our weakly-supervised HoverNet baseline is trained on pseudo nuclei masks generated using superpixels. We apply the SLIC [1] algorithm to partition the input image into superpixels. Each superpixel that contains an annotation point is considered the mask of a nucleus. In order to account for the variation in cell size and shape across different classes, we fine-tune the superpixel algorithm parameters in a class-specific manner.

In Figure 3, we visualize the superpixels for different classes of the Consep dataset. For reference, we also show the ground truth segmentation mask which was provided in Consep dataset. We observe that the superpixel-based pseudo-masks reasonably capture the scale and shape of the cells, but lose details along the boundary.

Faster-RCNN. For the baseline Faster-RCNN, we use the standard C4 architecture [4] with a ResNet101 feature encoder. All models were trained for 50,000 iterations. The only significant modification was in the maximal number of predicted instances, which we set to 1000 to allow for patches with a large number of cells, and the test ROI confidence threshold, which we set to 0.7 during inference.

Cascade-RCNN For the baseline Cascade-RCNN, we use the same encoder as in the paper [2], which is a Resnet101 based feature pyramid network. All models were trained for 50,000 iterations. To tune the model for this data domain, we set the anchor sizes to [8, 16, 32, 64, 128]. We find that the iterative process of detecting the bounding boxes used by Cascade-RCNN results in too many correct predictions failing the confidence threshold. As such, we both increase the maximum number of detections per image to 10000 and decrease the test ROI confidence threshold to 0.6.

PointSeg. The PointSeg baseline is based on the weakly-supervised segmentation model from [3]. We use this model to segment nuclei from an input patch. Next, we use the SSPP classification network [5] to perform classification. As in the original paper this network is trained on 27x27 20x magnification patches based on the point annotations. To account for variance in detected center, each point has up to 9 training patches selected, one centered at the point, and the other eight with centers up to 6 pixels away from the annotation point in varying directions. At inference stage, we use the weakly supervised segmentation network to predict segments of nuclei. For each segment, we use its centroid as the predicted cell location, extract a patch centered at the location, and apply the classification network to predict its class.

D. Average K-function Curves

In Fig. 4, the average pairwise K-curves are plotted for each pair of classes in each cluster. It shows how different clusters demonstrate different spatial behavior across pairs of cell classes. This also shows that the K-function across clusters is able to capture the variation in the spatial context.

E. Additional Experiments

We provide additional ablation studies on the hyper-parameters selection for the maximum radius of the K-function and the number of clusters in the deep clustering module.

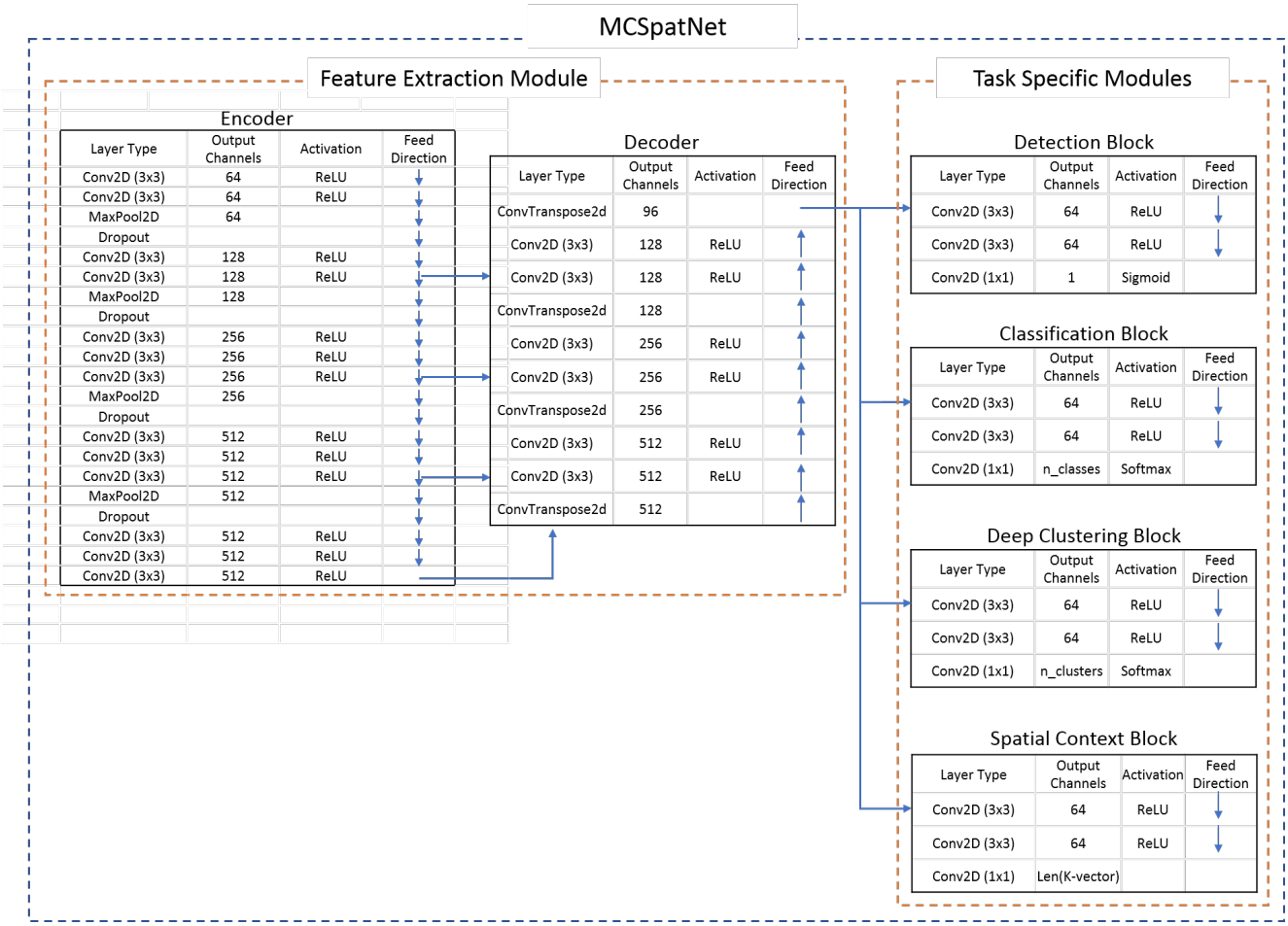

Figure 1. Model architecture details.

Table 1. Datasets statistics. For each dataset, we report number of training/validation/test patches. For each class, we also report numbers of cells in training/validation/test sets.

Dataset	N Patches	Inflam.	Epi.	Stroma
BRCA-M2C	80 / 10 / 30	3541 / 1358 / 960	9956 / 733 / 6109	5150 / 1042 / 1789
SEER-Lung	37 / 5 / 15	5499 / 726 / 3085	10875 / 1601 / 4906	5708 / 407 / 1871
Consep	22 / 5 / 14	2982 / 652 / 1537	4987 / 376 / 3113	4394 / 1343 / 3640

Varying K-function radius. Table 2 shows the results of varying the maximum radius of the K-function on the BRCA-M2C dataset. We experiment with radii of 60, 90, and 120 pixels. There is practically no difference in the detection F-score but there is a clear difference in the classification F-scores, which again shows the advantage of the spatial context learning to the classification task. 60 and 120 appear to be too small and too large radii. The best performance is at 90 pixels which is the setting we used in our experiments.

Varying number of clusters. Table 3 shows the results of varying the number of clusters in the deep clustering module. We experiment on the BRCA-M2C dataset with 3, 5, and 7 clusters. Again, the variation has no effect on the detection task. There are slight differences on the classification results with the best being at 5 clusters, which is the setting used in our experiments.

F. Additional Qualitative Results

We provide more qualitative examples to demonstrate the benefit of our proposed method, MCSpatNet, in Figures 5 and 6. Both figures show that MCSpatNet has better

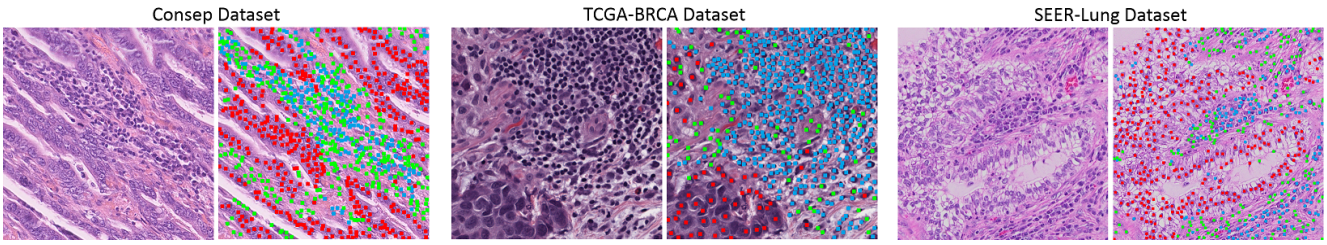

Figure 2. Sample patches and the corresponding point annotation. Colors indicate cell types: red=epithelial cells, blue=inflammatory cells, green=stromal cells.

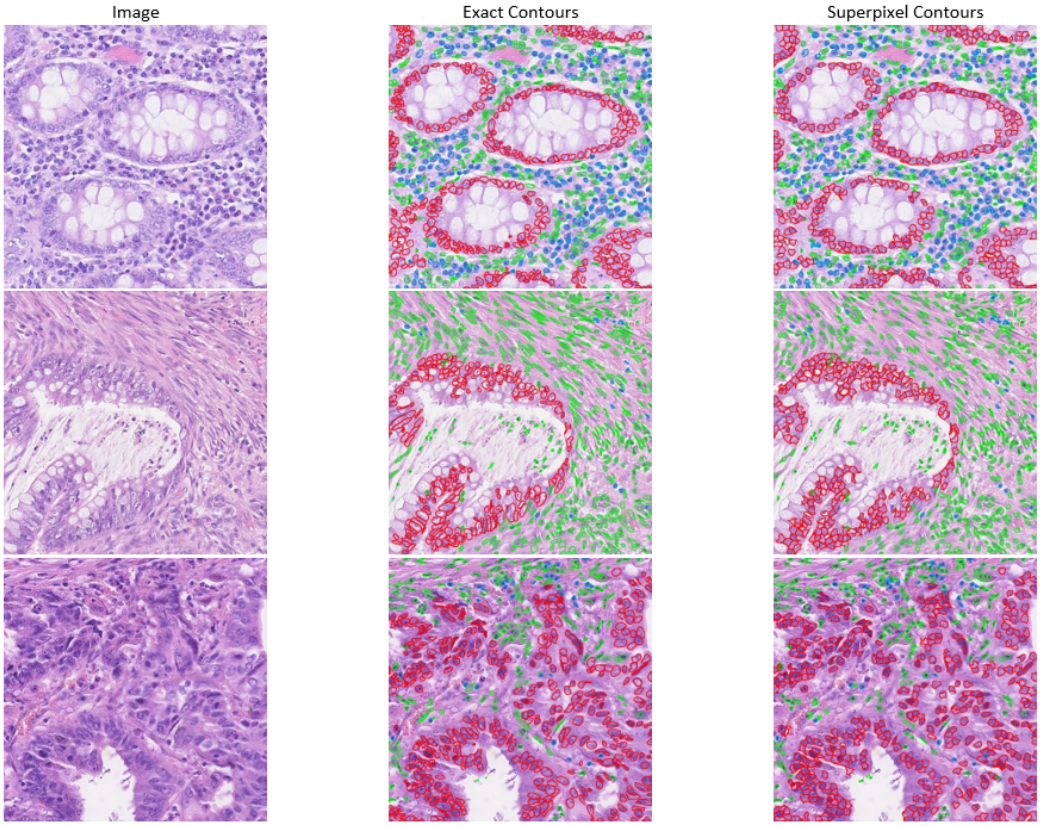

Figure 3. Sample images from Consep dataset, the exact segmentation masks (exact contours) provided by the original dataset, and superpixel-based pseudo masks (superpixel contours).

classification accuracy. Figure 5 demonstrates how MCSpatNet can better detect and classify cells even in very dense tumor nest regions. Figure 6 shows that MCSpatNet performs equally well in various tissue regions.

Table 2. Ablation study: Varying the K-function maximum radius on the BRCA-M2C dataset. Evaluation with F-scores per class, mean F-score over classes and detection F-score over all cells. Infl.: Inflammatory cells, Epi.: Epithelial cells, Stro.: Stromal cells, Mean: the mean F-score over the 3 classes, Det.: detection F-score.

Radius	Infl.	Epi.	Stro.	Mean	Det.
60	0.612	0.777	0.526	0.638	0.851
90	0.635	0.785	0.553	0.658	0.849
120	0.594	0.761	0.507	0.621	0.854

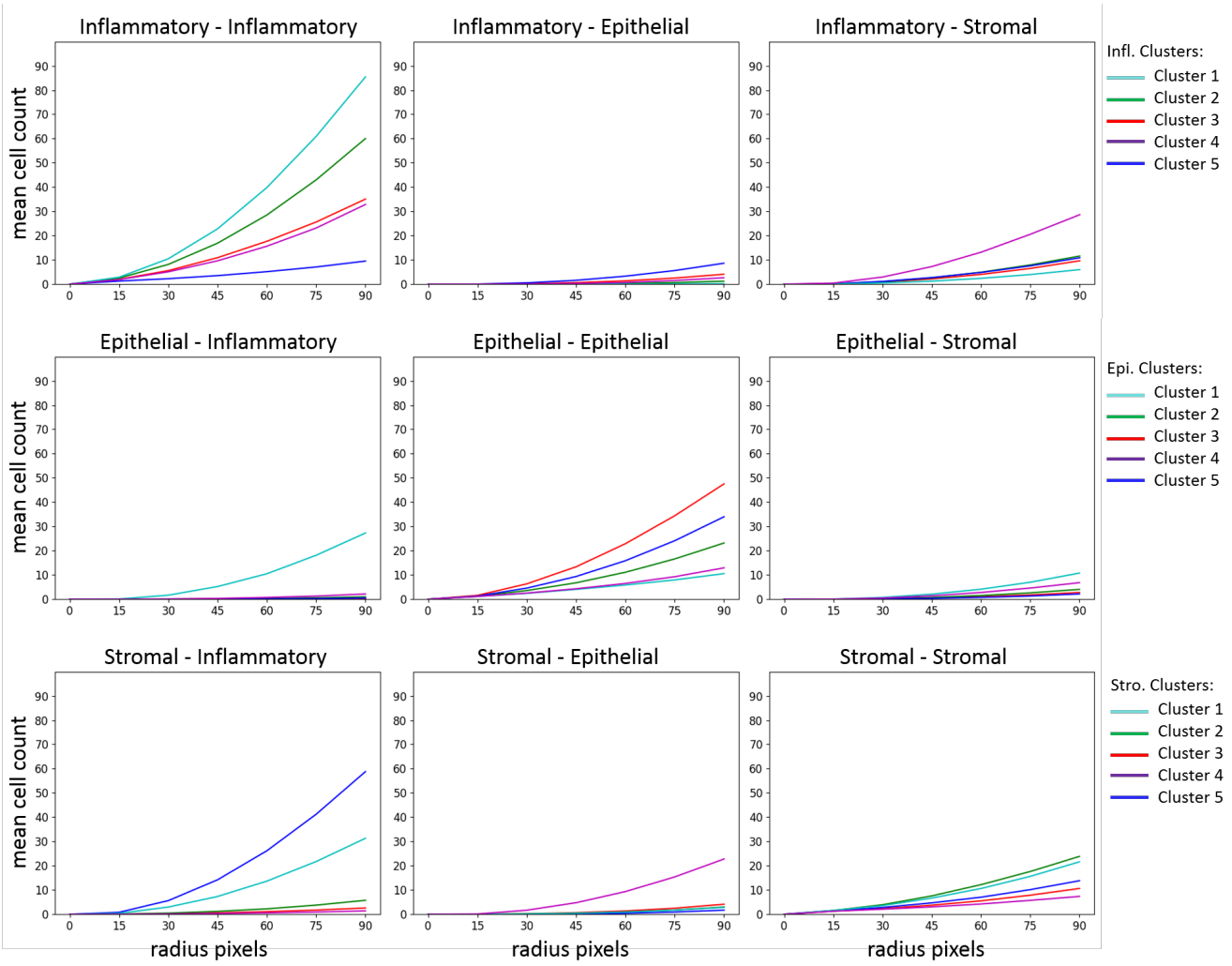

Figure 4. Average K-curves per cell type pair in each cluster. Each plot is for a pair of cell classes. The colored curves represent the K-curve for that pair in the different clusters. This clearly shows that different clusters demonstrate different cellular spatial behavior.

Table 3. Ablation study: Varying the number of clusters in the deep clustering module on the BRCA-M2C dataset. Evaluation with F-scores per class, mean F-score over classes and detection F-score over all cells. Infl.: Inflammatory cells, Epi.: Epithelial cells, Stro.: Stromal cells, Mean: the mean F-score over the 3 classes, Det.: detection F-score.

Clusters	Infl.	Epi.	Stro.	Mean	Det.
3	0.622	0.760	0.508	0.630	0.850
5	0.635	0.785	0.553	0.658	0.849
7	0.630	0.776	0.526	0.644	0.850

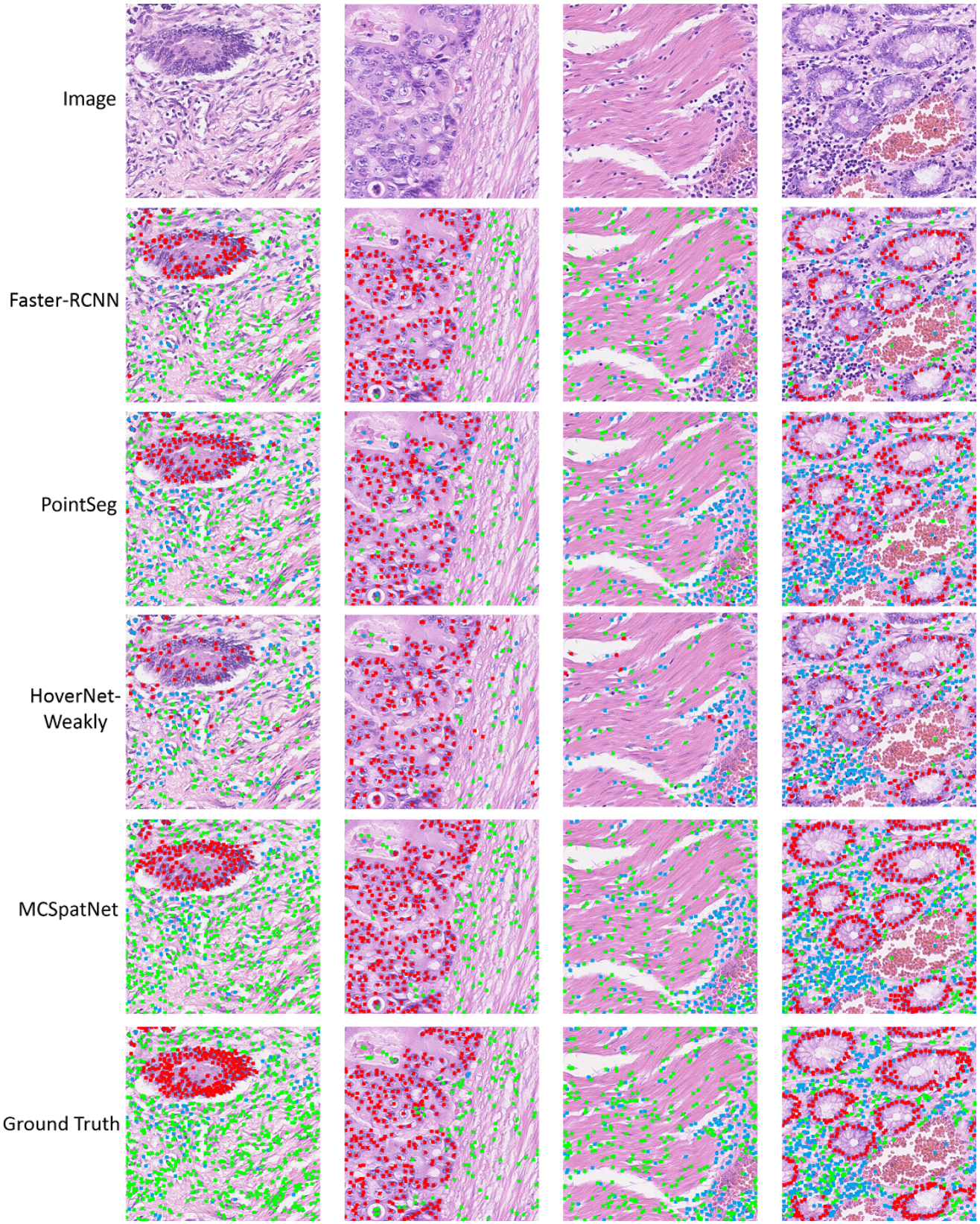

Figure 5. Qualitative results comparing against baseline methods in dense tumor regions. Top row are the original image patches. Bottom row are the ground truths. (Blue=inflammatory cells, Red=epithelial cells, Green=stromal cells.)

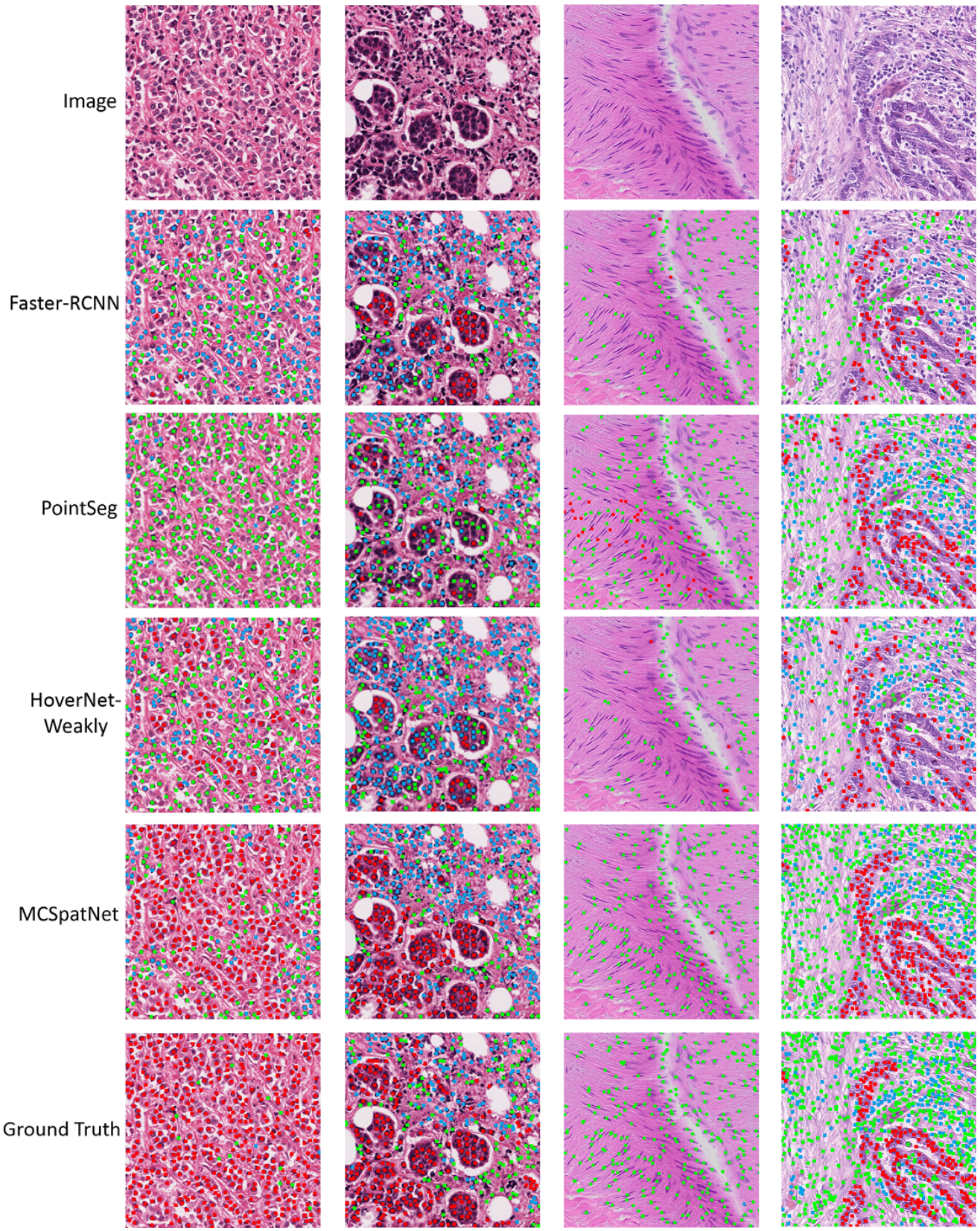

Figure 6. Qualitative results comparing against baseline methods in various tissue regions. Top row are the original image patches. Bottom row are the ground truths. (Blue=inflammatory cells, Red=epithelial cells, Green=stromal cells.)

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. 34(11):2274–2282, 2012. 1

[2] Zhaowei Cai and Nuno Vasconcelos. Cascade r-cnn: Delving into high quality object detection. 2018. 1

[3] Hui Qu, Pengxiang Wu, Qiaoying Huang, Jingru Yi, Gregory M. Riedlinger, Subhajyoti De, and Dimitris N. Metaxas. Weakly supervised deep nuclei segmentation using points annotation in histopathology images. 2019. 1

[4] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. 2015. 1

[5] K. Sirinukunwattana, S. E. A. Raza, Y. Tsang, D. R. J. Snead, I. A. Cree, and N. M. Rajpoot. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. 35(5):1196–1206, 2016. 1