Foveal Blur-Boosted Segmentation of Nuclei in Histopathology Images with Shape Prior Knowledge and Probability Map Constraints

HONGYI DUANMU [1,], FUSHENG WANG [1,2*], GEORGE TEODORO [3], AND JUN KONG [4,5*]

[1] *Department of Computer Science, Stony Brook University, NY, USA,* [2] *Department of Biomedical Informatics, Stony Brook University, NY, USA,* [3] *Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte, Brazil,* [4] *Department of Mathematics and Statistics and Computer Science, Georgia State University, Atlanta, GA,USA, and* [5] *Department of Computer Science and Winship Cancer Institute, Emory University, Atlanta, GA, USA.*

1. FOVEAL BLUR FOR BOOSTING DATASETS

As deep learning systems require a large scale of training images and annotations for effective training, we use foveal blur to rapidly enrich the dataset by generating 'nuclei-level' pathology image patches. In Fig. S1, we plot segmentation performances of different deep learning systems for comparison by Jaccard coefficient and accuracy when they are trained with different numbers of training images from the GBM dataset. Results in the other two datasets share similar patterns. Our system FB-Net trained with foveal blurred images achieves the best performance regardless of the training image number. In the most extreme case, we provide only one image patch for system training. In this case, our system outperforms other methods the most. As the number of training images increases, our system consistently outperforms other methods, but with a smaller advantage. It is noticed that all tested deep learning systems have limited generalizability to perform well over the testing set when they are trained with a small number of training images with limited diverse input information. With the help of foveal blur deriving a large number of nuclei-level images from one raw image, our system is able to yield 0.7021 by Jaccard coefficient when it is trained with only one image, only a 5% decrease from its performance associated with seven training images. DeepLabV3+ system achieves competitive results when it is trained with seven training images capturing about 1,000 nuclei. However, its performance drops dramatically to 0.6667 by Jaccard coefficient when trained with one image. By contrast, cGAN and U-Net both perform consistently worse in all cases and are more sensitive to the training data scale. Additionally, it is noticed that all methods for comparison present similar patterns by Accuracy. Thanks to the foveal blur boosting the training dataset, our system FB-Net with U-Net as its backbone is more robust and generalizable with a limited dataset than other methods for comparison.

2. EXPLICIT MORPHOLOGICAL KNOWLEDGE LEARNING BY MODIFIED LOSS FUNCTION

As demonstrated in the methodology section, the newly designed shape prior term in the loss function makes the deep learning model explicitly learn representative nuclei shapes and conforms its prediction to be compliant with the prior shape knowledge. Furthermore, the term of smoothness in the predicted probability map is designed to prevent the system from producing predictions with uneven edges.

In Fig. S2, the nuclei size histograms from our different deep learning models (i.e., FB-Net, FB+SP, and FB+SP+SM) are compared to the ground truth for the GBM dataset. In ground truth annotations, no nucleus is under 100 pixels in size. Note that more than 30% of nuclei produced by the FB-Net system are smaller than 100 pixels in size, while the FB+SP and FB+SP+SM systems reduce the percentage of such nuclei prediction to about 15% and 10%, respectively. The

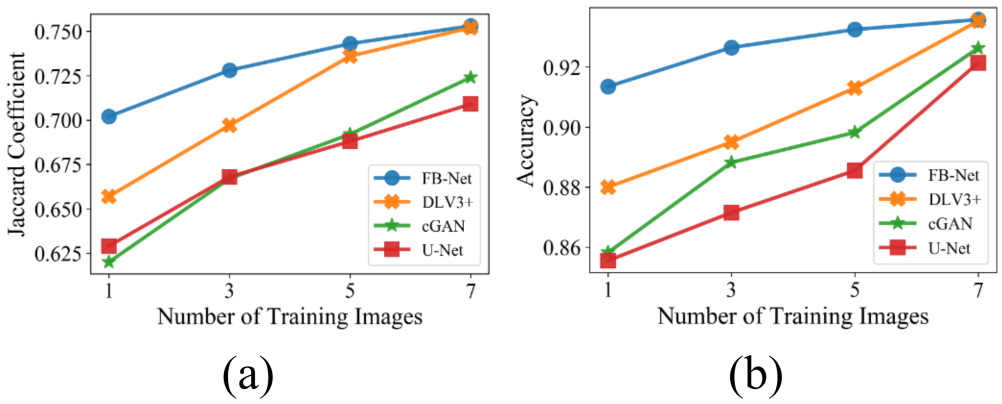

Fig. S1. Two performance plots are presented to compare (a) Jaccard coefficient and (b) Accuracy of diverse models trained with different numbers of images from the GBM dataset.

experimental results suggest that explicitly encapsulating the human knowledge, shape prior term and smoothness term can effectively restrain nuclei segmentation systems from producing unduly small nuclei.

We notice that the average nuclei size from the FB-Net, FB+SP, and FB+SP+SM systems are 326.4, 397.2, and 429.2 by pixels, respectively, while the average nuclei size from ground truth is 460.2 by pixels. The median and interquartile range (IQR) of nuclei size share a similar pattern. The medians of the nuclei size for FB-Net, FB+SP, FB+SP+SM, and ground truth are 313.0, 410.0, 442.0, and 448.5, respectively. The IQR of the nuclei size for FB-Net, FB+SP, FB+SP+SM, and ground truth are (60.0, 504.0), (209.5, 543.0), (262.0, 544.0), and (326.0, 569.0), respectively. These results suggest that the FB+SP and FB+SP+SM systems produce results more similar to the ground truth than the FB-Net system in terms of average nuclei size and two new loss terms improve the system's ability in generating nuclei segmentation with common size and shapes.

Fig. S3 demonstrates the performance improvement resulting from our two new loss function terms. Specifically, we present typical segmentation results of three regions of interest from our proposed three systems, i.e., FB-Net, FB+SP, and FB+SP+SM, respectively. Compared to the ground truth annotated by experienced pathologists, the FB-Net generates more small false positive nuclei regions than FB+SP and FB+SP+SM. These false positives can be readily avoided by human reviewers because of their irrational shapes and unduly small sizes. Thanks to the shape prior and morphology smoothness terms in the loss function, significantly less false positive nuclei are produced. In addition, it is noted that the smoothness term makes the prediction map smoother, making it easier for nuclei segregation during the post-processing step. By these experiments, we conclude human-knowledge boosted deep learning produces enhanced segmentation results.

In addition to results from our proposed system, we demonstrate typical segmentation results of conventional and some state-of-the-art deep learning methods for comparison in Fig. S4, and Fig. S5, respectively.

REFERENCES

1. M. Mirza and S. Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.
2. L. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818, 2018.
3. O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
4. H. Xu, C. Lu, R. Berendt, N. Jha, and M. Mandal. Automatic nuclear segmentation using multiscale radial line scanning with dynamic programming. *IEEE Transactions on Biomedical Engineering*, 64(10):2475–2485, 2017.
5. C. Jung and C. Kim. Segmenting clustered nuclei using h-minima transform-based marker

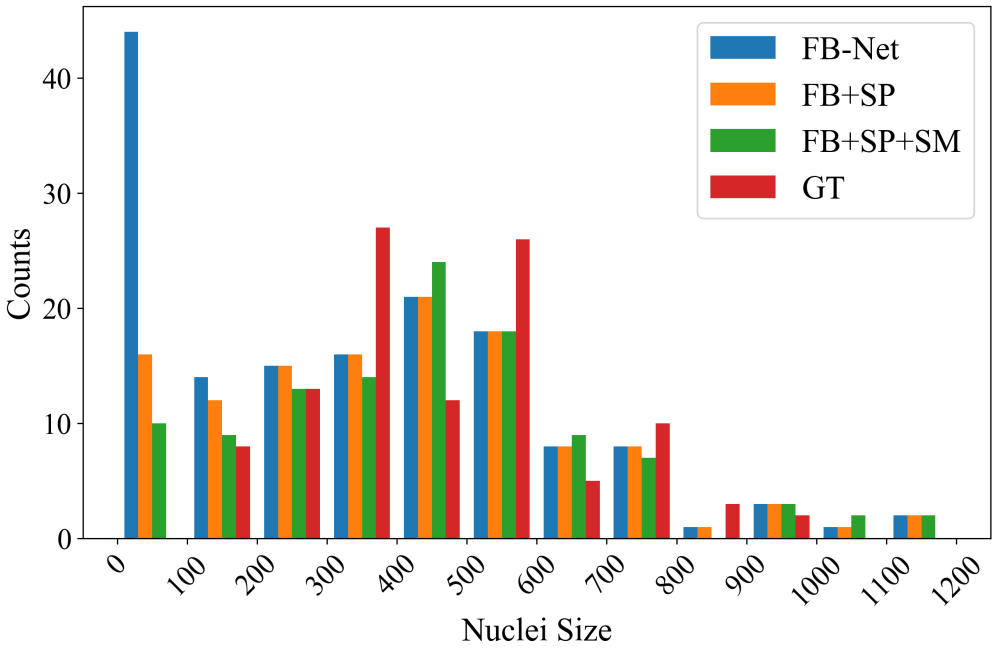

Fig. S2. Nuclei size histograms are presented with the ground truth from the GBM dataset and prediction results from our proposed models (i.e., FB-Net, FB+SP, and FB+SP+SM).

extraction and contour parameterization. *IEEE Transactions on Biomedical Engineering*, 57(10):2600–2604, 2010.

6. H. Xu, C. Lu, and M. Mandal. An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm. *IEEE Journal of Biomedical and Health Informatics*, 18(5):1729–1741, 2014.
7. F. Xing, H. Su, J. Neltner, and L. Yang. Automatic ki-67 counting using robust cell detection and online dictionary learning. *IEEE Transactions on Biomedical Engineering*, 61(3):859–870, 2014.

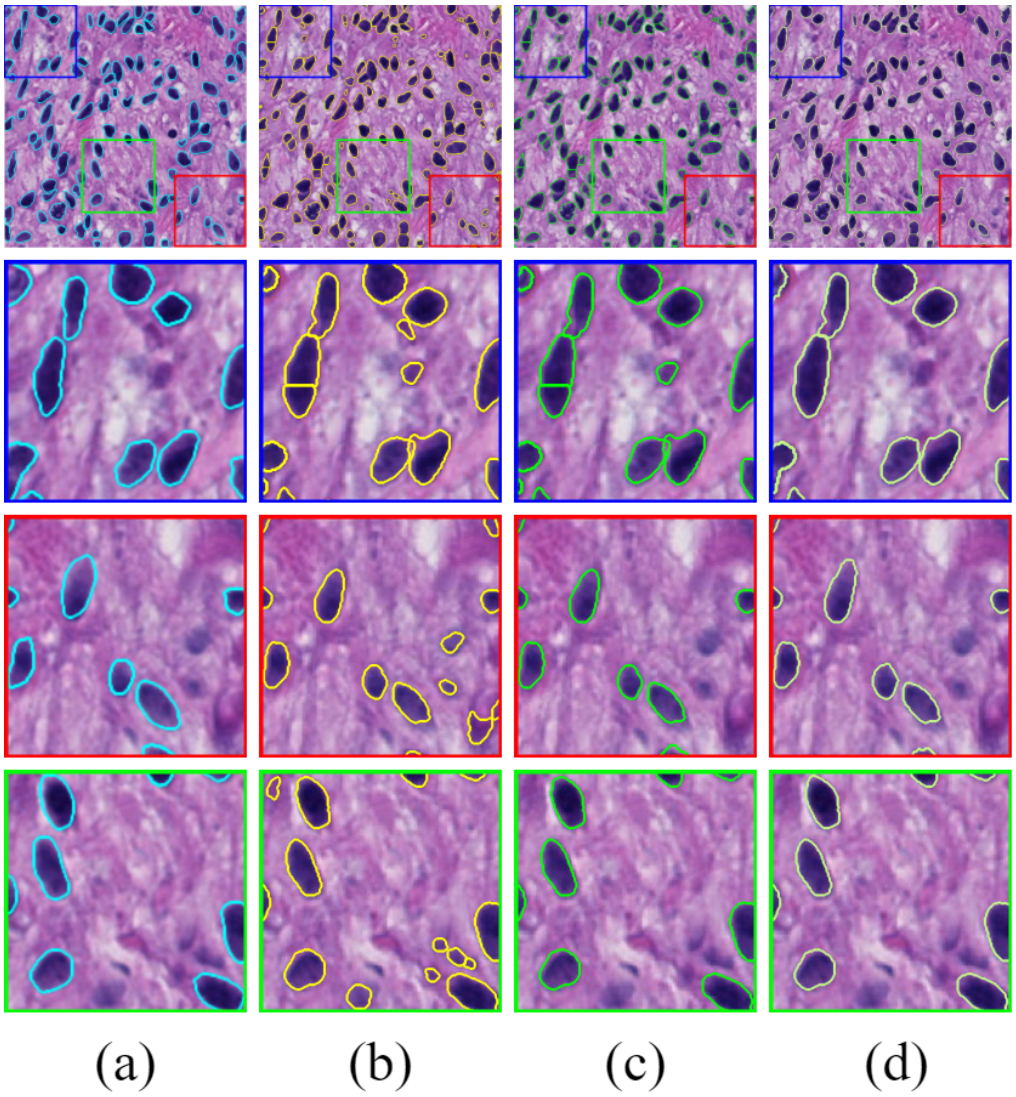

Fig. S3. We present a) ground truth, and representative segmentation results from b) FB-Net, c) FB+SP, and d) FB+SP+SM. It is noted that our new loss function terms help reduce small false positive nuclei and make the nuclei contour more regulated.

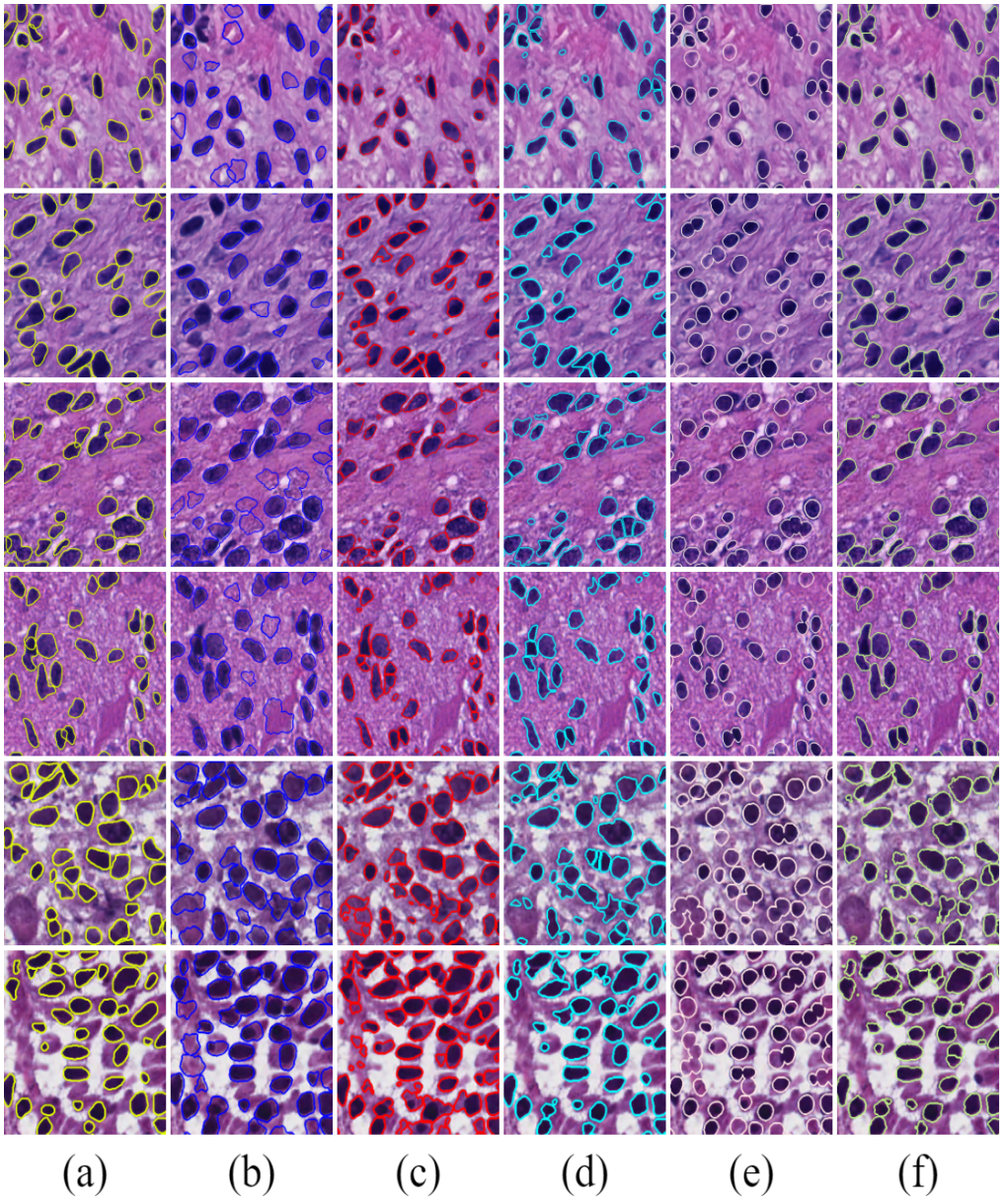

Fig. S4. We present typical segmentation results of conventional methods and our proposed system for comparison. The first, the second and the last two rows are histopathology images from GBM, TCGA, and MoNuSeg datasets, respectively. The columns from left to right present (a) ground truth, and results from (b) mRLS, (c) MOW, (d) IVW, (e) RACM, and (f) FB+SP+SM, respectively.

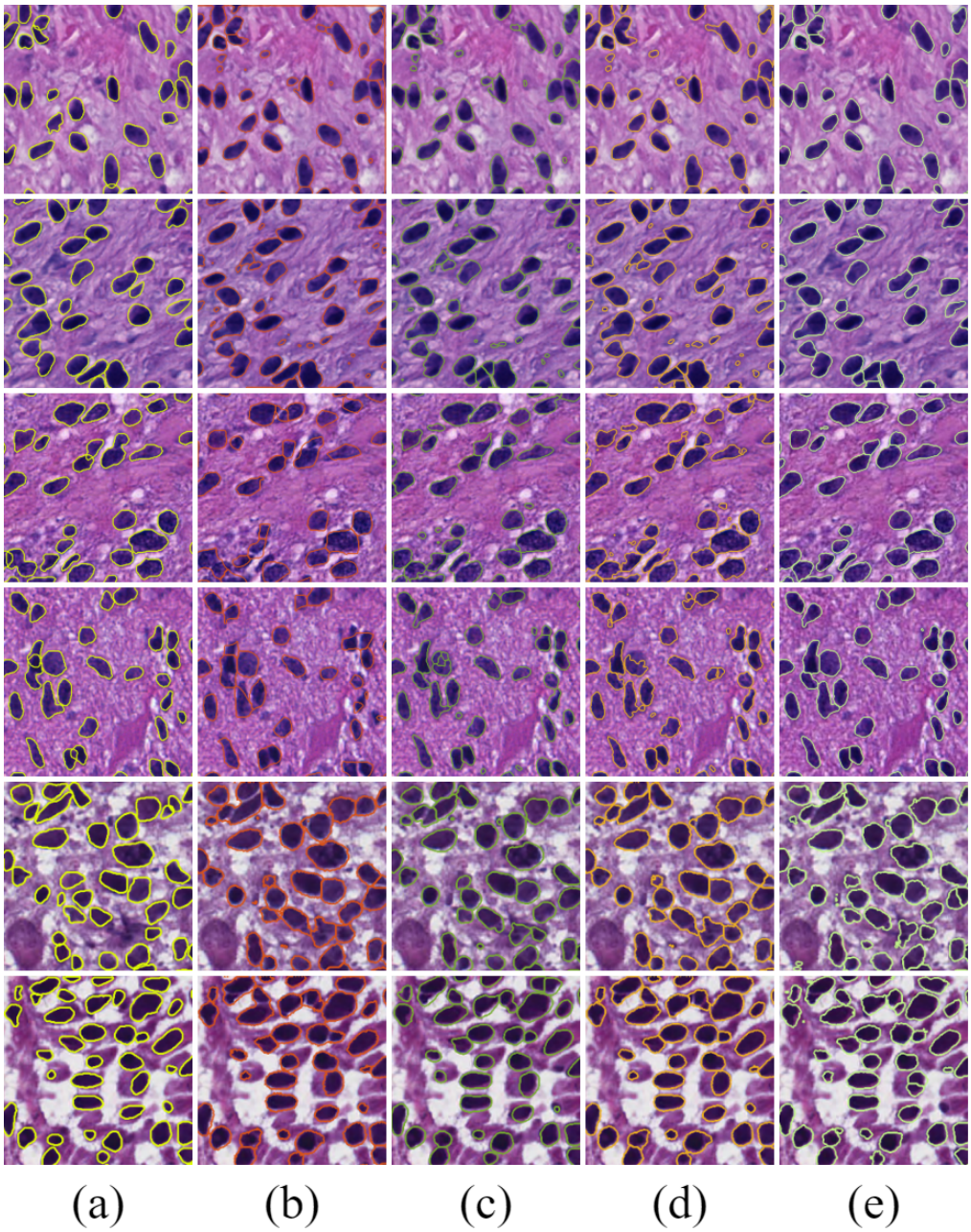

Fig. S5. We present typical segmentation results of some state-of-the-art deep learning methods and our proposed system for comparison. The first, the second, and the last two rows present histopathology images from GBM, TCGA, and MoNuSeg dataset, respectively. The columns from left to right present (a) ground truth, and results of (b) U-Net+CE, (c) DeepLabV3+, (d) cGAN, and (e) FB+SP+SM, respectively.