Robust Segmentation of Nucleus in Histopathology Images via Mask R-CNN

Xinpeng Xie[1], Yuexiang Li[2], Menglu Zhang[1], and Linlin Shen[1(✉)]

[1] Computer Vision Institute, Shenzhen University,
Shenzhen, Guangdong, China
{xiexinpeng2017,1810272064}@email.szu.edu.cn,
llshen@szu.edu.cn
[2] Youtu Lab, Tencent, Shenzhen, China
vicyxli@tencent.com

Abstract. Nuclei segmentation plays an import role in histopathology images analysis. Deep learning approaches have shown its strength for histopathology images processing in various studies. In this paper, we proposed a novel deep learning framework for automatic nuclei segmentation. The framework adopts the Mask R-CNN as backbone and employs structure-preserving color normalization (SPCN) and watershed for pre- and post-processing. The proposed framework achieved a Dice score of 90.46% on the validation set, which demonstrates its competing segmentation performance.

Keywords: Nuclei segmentation · SPCN · Deep learning · Instance segmentation

1 Introduction

Nuclei morphology plays an import role in identifying aberrant phenotypes, which needs to detect and segment all the nuclear in hematoxylin and eosin (H&E) stained histology images. The features and distribution of nucleus can provide direct and reliable information for the diagnosis of cancer. The main challenges of nuclei segmentation are: (i) wide variation of cell appearance, such as color, shape, and texture; (ii) weak/missing cell boundaries; (iii) crowding and overlapping cells. With the development of deep learning technique, it is worthwhile to design and implement more accurate and efficient algorithms for nuclei segmentation. In [1], Bengtsson et al. proposed a seeded watershed transform to incorporate intensity, gradient, connectivity, and shape information for nucleus segmentation. Classical techniques are most procedural and require a large number of free parameters, Ronneberger et al. pretented the U-net, a novel FCN based network architecture, for biomedical image segmentation [2] and won the Grand Challenge for Computer-Automated Detection of Caries in Bitewing Radiography at ISBI 2015. In [3], Cui et al. designed a deep learning algorithm for one-step contour aware nuclei segmentation. They introduced a nucleus-boundary model to predict nuclei and boundaries simultaneously. Khoshdeli et al. [4] integrated boundary- and region-based information to distinguish touching or overlapping nuclei. They labeled foreground (nuclei), background and boundary pixels, to

train region- and boundary-based models, respectively. Then they trained another convolutional neural network to fuse the information from the two models to get the final result.

In this work, we propose a deep learning based framework for nuclei segmentation of Glioma whole slide tissue images. The main contributions of this work are listed in the following:

(1) We adopt structure-pre-serving color normalization (SPCN) [5] for sparse stain separation and color normalization, which can significantly boost the performance of deep learning framwork.
(2) We develop an effective framework based on Mask R-CNN to address the problem of overlapping nuclus segmentating, which achieves competitive results on MICCAI 2018 competition: Segmentation of Nuclei in Images.

2   Method

2.1   Dataset

The Digital Pathology dataset in [6] is adopted in this study. The dataset contains 15 pieces of about 668×583 annotated H&E stain images extracted from a set of Glioblastoma and Lower Grade Glioma whole slide tissue images. We separate the dataset to training and validation sets according to the ratio of 80:20. As the dataset is too small to train a deep learning network directly, two extra datasets, i.e. MICCAI 2017 segmentation challenge (MSC) [7] and 2018 Data Science Bowl (DSB) [8], are used for data augmentation. The MSC dataset includes 32 annotated H&E stain images from four categories, i.e. hnsc, lung, lgg and gbm. Each category has 8 images. The DSB dataset contains 664 fluorescent stain and H&E stain images. To alleviate color variation, we remove the fluorescent stain images in DSB dataset, which results in a dataset with 108 images.

2.2   Pre-processing

To overcome the color variation, structure-pre-serving color normalization (SPCN) [5] is adopted in our framework. For a source image S and a target image T, their color appearances and stain density maps is first estimated by factorizing V_s into $W_s H_s$, and V_t into $W_t H_t$ using sparse non-negative matrix factorization. Then, combined with color appearance of the target W_t, a scaled version of the density map of source H_s is calculated to generate the normalized source image, which can be described as follows:

$$H_s^{norm}(j,:) = \frac{H_s(j,:)}{H_s^{RM}(j,:)} H_t^{RM}(j,:), j = 1, \ldots, r.$$

where V_s, W_s, H_s repensent the observation matrix, stain color appearance matrix and stain density map matrix of source image s, respectively. $H_i^{RM} = RM(H_i) \in R^{r \times 1}, i = (s,t)$ and $RM(\cdot)$ compute robust pseudo maximum of each row vector at 99%.

In our experiment, we noticed that the E stain image contains tissue information, while the cell nucleus information is mainly contained in the H stain images. Therefore, we use the H stain images for network training. Some separated H&E stain images are presented in Fig. 1.
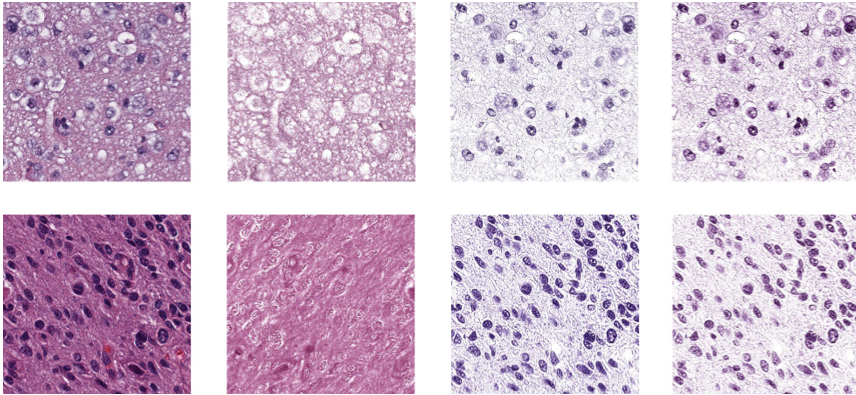
Fig. 1. Some examples of sparse stain separation for histological images. The first column is the original histological images, the second and third column are the E stain, H stain generated by SNMF, respectively. The last column is the H stain image after color normalization.

2.3   The Proposed Segmentation Framework

The flowchat of our segmentation framework is shown in Fig. 2. The Mask R-CNN [9] is employed as the backbone of our model. First, SPCN [5] is adopted to process original images. Each nuclei is separated from ground truth for RoI extraction. We slide a 512×512 window to crop patches from the whole slices and train the Mask R-CNN. We employ resnet101 as the feature extraction network. The watershed algorithm is employed as the post-processing to separate the joint cells. ImageNet [10] and COCO [11] datasets are adopted for network pre-training.
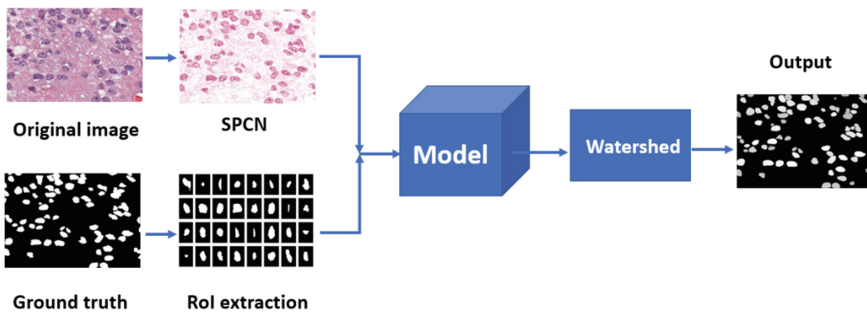
**Fig. 2.** The framework of our segmentation model.

Mask R-CNN. Figure 3 shows the Mask R-CNN [9] framework for nuclei segmentation, which has multiple architectures, i.e. (i) the convolutional backbone architecture for feature extraction, (ii) the network head for bounding-box recognition and mask prediction. The RoIAlign layer is developed to address misalignments between the RoI and the extracted features caused by RoIPool [12], which is simple but useful to predict pixel-level masks.
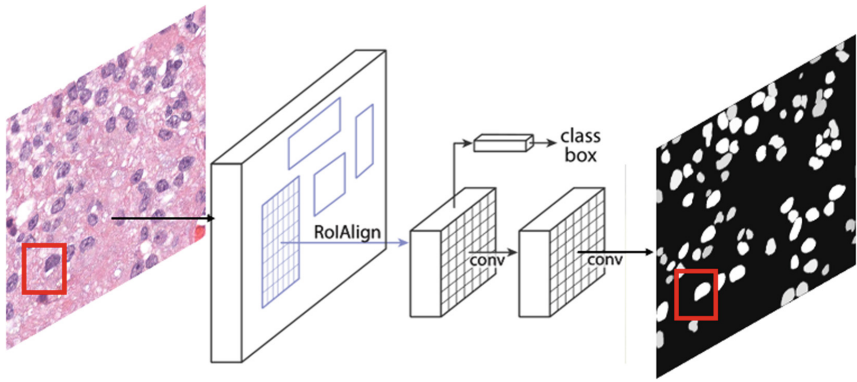
**Fig. 3.** The Mask R-CNN framework for nuclei segmentation.

2.4   Post-processing

Combining the information of detection, classification and bounding-box regression, MaskR-CNN [9] can efficiently segment each nuclei in images. However, not all of the overlapping nuclei are accurately separated. An example is shown in Fig. 4. Therefore, to address the problem, the Marker-based watershed [13] is involved in our framework. Given two nuclei S and G, their overlapping ratio $P_s = \frac{O}{S}$ and $P_g = \frac{O}{G}$, where O
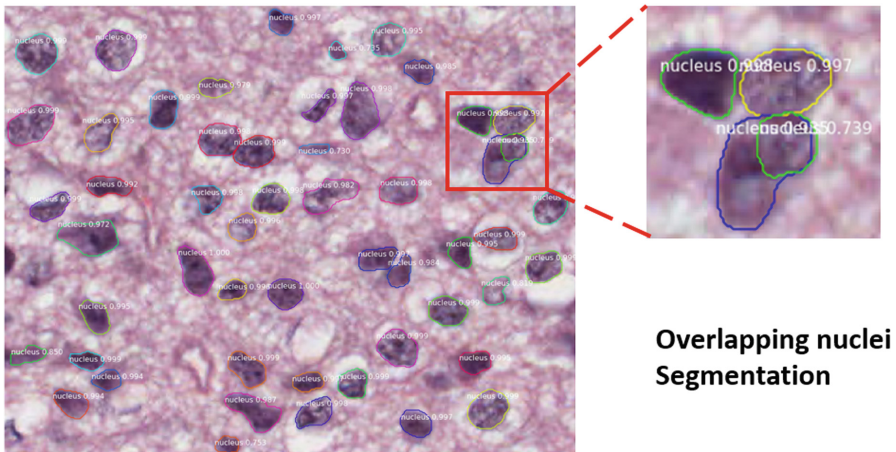
Fig. 4. An example of overlapping nuclei segmentation.

represents the overlapping area of S and G, are calculated to decide the overlapping area of each nuclei. The watershed post-processing result is shown in Fig. 5 (Better seen in color).
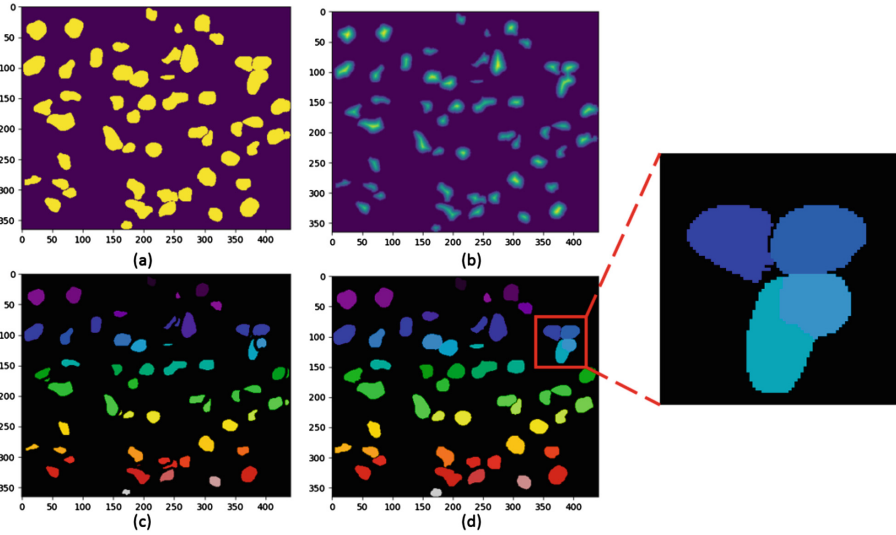
Fig. 5. Marker-based watershed post-processing. (a) is the original input mask; (b) is the marker from the minima computed from the distance transform of the binarized mask; (c) is the marker after binary_openning; (d) is the final output.

2.5 Implementation

The proposed framework is implemented using Keras toolbox, and trained with a mini-batch size of 64 on four GPUs (GeForce GTX TITAN X, 12 GB RAM). The initial learning rate is set to 0.0001. 'Adam' [14] is used to iteratively update neural network weights based on training data.

3 Experimental Results

3.1 Evaluation Criterion

The Dice score, a measure for comparison of binary segmentations S and G, is introduced to evaluate the performances on the validation set. It can be expressed in terms of statistical measures as:

$$D = \frac{2|S \cap G|}{|S| + |G|} = \frac{2\theta_{TP}}{2\theta_{TP} + \theta_{FP} + \theta_{FN}}$$

where θ_{TP} is the number of true positives, θ_{FP}/θ_{FN} are the numbers of false positives/false negatives.

3.2   Results

We first test the performance of our network trained using different training data. The dice scores on validation set trained with different datasets are summarized in Table 1. MSC and DSB represent the dataset from the MICCAI 2017 segmentation challenge [7] and the 2018 Data Science Bowl [8], respectively. Fixed DSB is the refined data after deleting fluorescent stain images from DSB. Compared with the network trained using the original Digital Pathology dataset [6], the network trained with MSC improves the dice score from 79.65% to 81.32%. While similar improvement was achieved for fixed DSB, the DSB actually decrease the dice score, which proves the necessity to clean fluorescent stain images in DSB. Trained with the combination of original dataset, MSC and fixed DSB, the Mask R-CNN achieves the highest dice score, i.e. 85.01%.

Table 1. Dice score for different training data (%).

Training data	Dice score
Digital Pathology	80.16
MSC	81.32
DSB	80.09
fixed DSB	82.15
Digital Pathology + MSC + fixed DSB	**85.01**

Once the training dataset is fixed, we now test the performance of the Mask R-CNN, pretrained with different dataset and the effect of preprocessing and postprocessing. As listed in Table 2, the model pre-trained on COCO dataset performances better than that pre-trained on ImageNet dataset, i.e. a 1.44% dice score improvement was recorded. Furthermore, SPCN [5] and Watershed [13] improve the Dice score on validation set by 1.04% and 0.62%, respectively. Combined with SPCN [5] and Watershed [13], Mask R-CNN achieves the highest dice score with 90.46%.

Table 2. Dice score for different processing (%).

Model	Pre-training on	Dice score (%)
Mask R-CNN	ImageNet	86.99
Mask R-CNN	COCO	88.43
Mask R-CNN + SPCN	COCO	89.47
Mask R-CNN + Watershed	COCO	89.05
Mask R-CNN + SPCN + Watershed	COCO	**90.46**
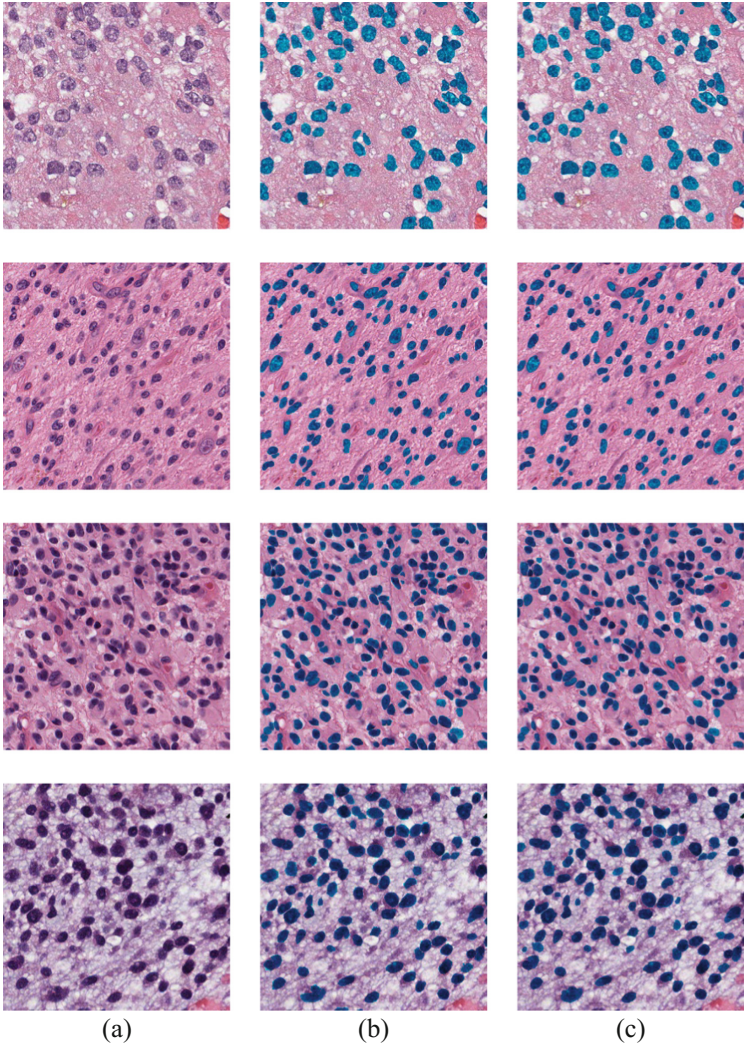
Fig. 6. Qualitative segmentation result on the data set. (a) The original input image (b) Visualizations of the ground truth (c) Visualizations of our model.

To visually assess our segmentation results, we overlapped the segmentation results and ground truths back to original images, as shown in Fig. 6.

CPM Competition Result. We achieved a competitive dice score of 86.10% on the MICCAI 2018 CPM competition testing set, while the highest score is 87.00%. Table 3 listed the Top-5 results on the leader-board.

Table 3. Dice scores for Top 5 (%).

Model	Dice score (%)
Top 1	87.00
Top 2	86.80
Top 3	86.20
Top 4 (ours)	**86.10**
Top 5	84.80

4   Conclusions

Diversity of phenotypes and overlapping nuclei are the two intrinsic challenges for nuclear segmentation in H&E stained images. In this paper, we proposed a Mask R-CNN based deep learning framework for nuclei segmentation. The overall framework consists of color normalization using SPCN [5], features extraction, and Watershed [13] postprocessing. The experimental results show that SPCN [5] and Watershed [13] can significantly improve segmentation performance, and our framework has strong strength in nuclei segmentation.

References

1. Bengtsson, E., Wahlby, C., Lindblad, J.: Robust cell image segmentation methods. Pattern Recogn. Image Anal. **14**(2), 157–167 (2004). C/c of Raspoznavaniye Obrazov i Analiz Izobrazhenii
2. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
3. Cui, Y., Zhang, G., Liu, Z., et al.: A deep learning algorithm for one-step contour aware nuclei segmentation of histopathological images. arXiv preprint arXiv:1803.02786 (2018)
4. Khoshdeli, M., Parvin, B.: Deep leaning models delineates multiple nuclear phenotypes in H&E stained histology sections. arXiv preprint arXiv:1802.04427 (2018)
5. Vahadane, A., et al.: Structure-preserving color normalization and sparse stain separation for histological images. IEEE Trans. Med. Imaging **35**(8), 1962–1971 (2016)
6. MICCAI CPM: Digital pathology: segmentation of nuclei in images. https://miccai.cloudapp.net/competitions/83
7. MICCAI 2017 segmentation challenge. https://miccai.cloudapp.net/competitions
8. Kaggle Data Science Bowl. https://www.kaggle.com/c/data-science-bowl-2018
9. He, K., Gkioxari, G., Dollar, P., Girshick, R.: Mask R-CNN. In: International Conference on Computer Vision (ICCV) (2017)

10. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: a large scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255 (2009)
11. Lin, T.-Y., et al.: Microsoft COCO: common objects in context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8693, pp. 740–755. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10602-1_48
12. Girshick, R.: Fast R-CNN. In: International Conference on Computer Vision (ICCV) (2015)
13. Dundar, M.M., et al.: Computerized classification of intraductal breast lesions using histopathological images. IEEE Trans. Biomed. Eng. **58**(7), 1977–1984 (2011)
14. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv e-print arXiv:1412. 6980 (2014)