



A Deep Model with Shape-Preserving Loss for Gland Instance Segmentation

Zengqiang Yan¹(✉), Xin Yang², and Kwang-Ting Tim Cheng¹

¹ Department of Computer Science and Engineering,
The Hong Kong University of Science and Technology, Kowloon, Hong Kong
z.yan@connect.ust.hk

² School of Electronic Information and Communications,
Huazhong University of Science and Technology, Wuhan, China

Abstract. Segmenting gland instance in histology images requires not only separating glands from a complex background but also identifying each gland individually via accurate boundary detection. This is a very challenging task due to lots of noises from the background, tiny gaps between adjacent glands, and the “coalescence” problem arising from adhesive gland instances. State-of-the-art methods adopted multi-channel/multi-task deep models to separately accomplish pixel-wise gland segmentation and boundary detection, yielding a high model complexity and difficulties in training. In this paper, we present a unified deep model with a new shape-preserving loss which facilities the training for both pixel-wise gland segmentation and boundary detection simultaneously. The proposed shape-preserving loss helps significantly reduce the model complexity and make the training process more controllable. Compared with the current state-of-the-art methods, the proposed deep model with the shape-preserving loss achieves the best overall performance on the 2015 MICCAI Gland Challenge dataset. In addition, the flexibility of integrating the proposed shape-preserving loss into any learning based medical image segmentation networks offers great potential for further performance improvement of other applications.

Keywords: Deep convolutional neural network
Gland instance segmentation · Shape-preserving loss
Histology image analysis

1 Introduction

Accurate gland instance segmentation from histology images is a crucial step for pathologists to quantitatively analyze the malignancy degree of adenocarcinomas for further diagnosis [1]. However, manual annotation of gland instances is time consuming given the large size of a histology image and a large number of glands in an image. Therefore, accurate and automatic methods for gland instance segmentation are in great demand. Gland instance segmentation consists of two

sub-tasks: (1) gland segmentation which separates glands from the background and (2) boundary detection so as to identify each gland individually, as shown in Fig. 1. In practice, gland instance segmentation is a challenging task due to the following two unique characteristics. First, the appearances of glands vary significantly in different histology image patches (or even within a single patch), as shown in Fig. 1. Large appearance variations make the pixel-wise gland segmentation problem very challenging. Second, some gland instances are very close to each other or even share one entity borders, making it hard to identify gland instances individually and preserve the shape of each gland.



Fig. 1. Gland instance segmentation problem: (a) histology image patch, (b) gland segmentation, (c) boundary detection and (d) ground-truth gland instance segmentation (different colors represent different gland instances). Gland instance segmentation can be regarded as the combination of boundary detection and gland segmentation.

Several methods have been proposed for accurate gland instance segmentation. Xu *et al.* [2, 3] proposed multi-channel deep neural networks to extract gland region, boundaries and location cues separately. The results from different channels were fused to produce the final gland instance segmentation result. In their methods, different channels and the fusion module were implemented using different deep learning architectures and were pre-trained separately. While achieving the best performance so far on the 2015 MICCAI Gland Challenge dataset, the method incurs a high model complexity and complicated training process. Chen *et al.* [4] formulated the gland instance segmentation problem as a multi-task learning problem. The model contains two branches trained using the manual annotations for gland segmentation and the manually generated boundary maps for boundary detection separately.

Different from the complicated multi-channel/-task deep models, we propose a shape-preserving loss which enables a single deep model to jointly learn pixel-wise gland segmentation and boundary detection simultaneously. The proposed shape-preserving loss mainly consists of two parts: (1) a pixel-wise cross-entropy loss for gland segmentation and (2) a shape similarity loss for gland boundary detection. Experimental results demonstrate that the deep model trained by the shape-preserving loss outperforms the state-of-the-art methods on the public gland segmentation dataset. In addition, the great flexibility of the shape-preserving loss for integration into any deep learning models is a significant benefit.

2 Method

A straightforward solution to learn a unified deep model for both gland segmentation and boundary detection is to minimize pixel-wise cross-entropy losses between: (1) detected gland pixels and annotated gland pixels (gland segmentation), (2) detected gland boundaries and annotated boundaries (gland boundary detection). However, learning a boundary detector via pixel-wise cross-entropy loss utilizes only information of individual pixels with little consideration of global shape information of an entire boundary segment. In addition, it is often very challenging to precisely localize gland boundary pixels due to the ambiguity at gland boundaries as well as low contrast, as shown in Fig. 2(a). Consequently, directly using the pixel-wise cross-entropy loss for gland instance segmentation would suffer from poor performance on shape preservation. Increasing the importance of boundary pixels could partially solve the problem while it would also degrade the gland segmentation performance due to numerous holes in gland segmentation results. Therefore, it is difficult for the pixel-wise cross-entropy loss to balance the importance between boundary detection and gland segmentation in the training process. In this section, we first present a novel shape-preserving loss to enable a unified deep model for both accurate and efficient boundary detection and gland segmentation followed by the corresponding deep model.

2.1 Shape-Preserving Loss

The proposed shape-preserving loss consists of two losses: (1) a shape similarity loss [5] which is able to detect boundary segments with similar geometric shapes as manual annotations while allows certain boundary location variations arising from low intensity contrast and ambiguities of boundaries; and (2) a weighted pixel-wise cross-entropy loss for close and/or adhesive gland segmentation. In the following, we describe each loss in more details.

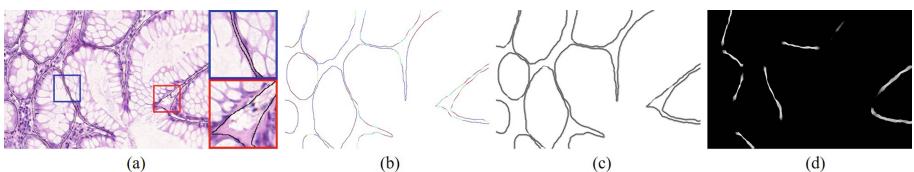


Fig. 2. Shape-preserving loss construction: (a) histology patch by remapping the annotated boundaries, (b) boundary segmentation map, (c) searching range map and (d) weight map background pixels between close glands. In the boundary segmentation map, different boundary segments are assigned with different colors. In the weight map, the intensity of each pixel represents its weight (0~1.0).

Shape Similarity Loss. Given a manual annotation, we first generate the corresponding boundary map and segment the whole boundary map into boundary segments (by default the length is in the range of [16, 24] pixels) as shown in Fig. 2(b). For each boundary segment l_i , as shown in Fig. 2(c) a 3-pixel searching range is assigned to find the corresponding boundary segment s_i in the output segmentation map. As the 3-pixel searching range allows limited location variation between l_i and s_i , the measurement is less restrictive and more efficient compared to the pixel-wise cross-entropy measurement. Then, the shape similarity is measured based on the curve fitting method. Given l_i and s_i , we adopt a curve (cubic function) fitting method to obtain their approximated curve functions $f(l_i) = a_1x^3 + b_1x^2 + c_1x + d_1$ and $f(s_i) = a_2x^3 + b_2x^2 + c_2x + d_2$. The shape similarity between l_i and s_i is defined as

$$ss_i = \left| \frac{\langle F_1, F_2 \rangle}{\|F_1\| \|F_2\|} \right| \in [0, 1], \quad (1)$$

where $F_1 = [a_1, b_1, c_1]$, $F_2 = [a_2, b_2, c_2]$, $\langle \cdot, \cdot \rangle$ is the dot product operation and $\|\cdot\|$ measures the length of the vector. Furthermore, if $|s_i| < 0.9 \cdot |l_i|$ or $|s_i| > 1.2 \cdot |l_i|$, the shape similarity ss_i is manually set as 0. Based on the shape similarity measurement, the shape similarity loss L_s for pixels in the generated segmentation map located within the 3-pixel searching range is constructed as $L_s = w_s \cdot (1 - ss)$, where w_s is set to 4 to increase the importance of the shape similarity loss in the overall loss calculation.

Weighted Pixel-Wise Cross-Entropy Loss. To better identify close and/or adhesive gland instances individually and inspired by the weight matrix in the *U-Net* model [6], we construct a weight matrix W which effectively increases the importance of those background pixels between close gland instances. For each background pixel i , the weight $w(i)$ is defined as

$$w(i) = 1 + w_0 \cdot \frac{1}{1 + e^{d_1(i) + d_2(i) - \mu}}, \quad (2)$$

where d_1 denotes the distance to the nearest gland and d_2 is the distance to the second nearest gland. In our experiments we set $w_0 = 6$ and $\mu = 15$ pixels.

Accordingly, the shape-preserving loss L can be formulated as

$$L = W \cdot L_p \cdot (1 + L_s), \quad (3)$$

where L_p is the overall pixel-wise cross-entropy loss for gland segmentation.

2.2 Deep Convolutional Neural Network

In terms of the deep learning architecture, we design the deep model with reference to the *HED* model [7,8] which was proposed for edge/contour detection. Since the deep model has multiple side outputs, the model is able to learn accurate boundary detection by the first several side outputs and accurate pixel-wise gland segmentation by the last few side outputs. As the shape-preserving loss re-balances the importance between boundary detection and pixel-wise gland segmentation, the training process is controllable and is able to converge quickly.

3 Evaluation and Discussion

We evaluated the proposed deep model on the 2015 MICCAI Gland Challenge dataset [9]. Three officially provided metrics are used for the performance evaluation: (1) the F1 score, (2) the object-level dice score and (3) the object-level Hausdorff distance. We compared our method with three state-of-the-art methods, *i.e.* Chen *et al.* [4] which achieved the best results in the 2015 MICCAI Gland Segmentation Challenge and the multi-channel deep models proposed by Xu *et al.* [2,3] which reported the best performance so far on this dataset. Table 1 shows the comparison results among these methods. Based on the construction of the proposed shape-preserving loss, the deep model is able to identify different glands individually which helps achieve the best performance in the F1 score. As the shape similarity measurement in the shape-preserving loss allows limited boundary location variation, the performance of the deep model in terms of the object-level dice score would be slightly influenced. Since the radius of the searching range adopted is set to only 3 pixels, the influence would be relatively limited and the deep model can still achieve top results. Similarly, as the shape similarity measurement is different from the definition of the Hausdorff distance, the performance may also be influenced. In short, the deep model with the proposed shape-preserving loss achieves the best overall performance (Fig. 3).

Exemplar segmentation results are shown in Fig. 4. Compared to the manual annotations, the deep model can effectively segment different glands individually and preserve shapes of glands accurately. Although boundaries of the segmented glands are slightly different from those annotated boundaries at the pixel-wise

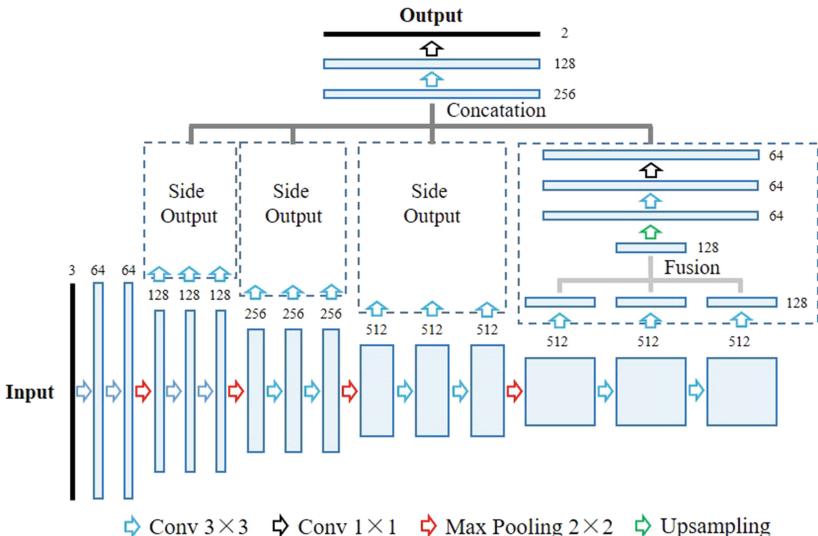


Fig. 3. The overview of the proposed deep model trained by the shape-preserving loss.

Table 1. Comparison results among different methods on the public dataset.

Method	F1 Score		ObjectDice		ObjectHausdorff	
	Part A	Part B	Part A	Part B	Part A	Part B
Ours	0.924	0.844	0.902	0.840	49.881	106.075
Xu [2]	0.893	0.843	0.908	0.833	44.129	116.821
Xu [3]	0.858	0.771	0.888	0.815	54.202	129.930
Chen [4]	0.912	0.716	0.897	0.781	45.418	160.347

**Fig. 4.** Exemplar segmentation results obtained by the proposed deep model. For each row, from left to right: raw histology patch, manual annotation, generated probability map and the output segmentation map. In both manual annotation and output segmentation map, different glands are assigned with different colors.

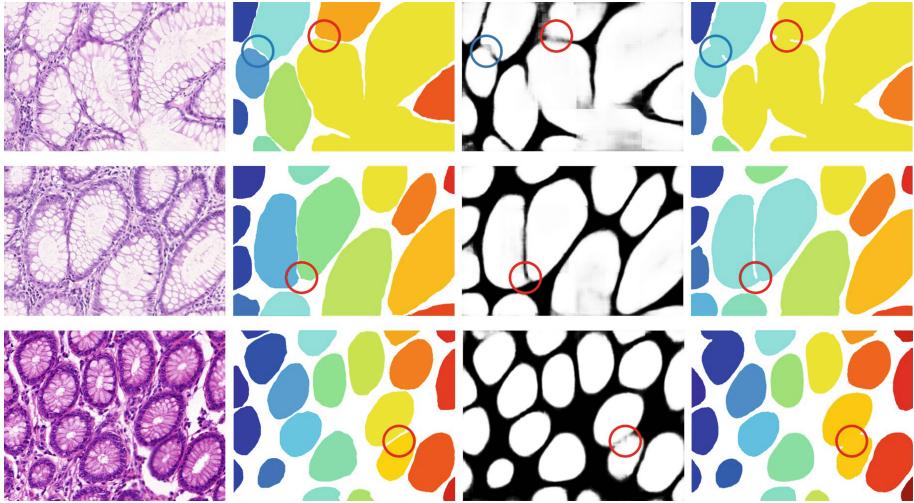


Fig. 5. Challenging cases in dealing with close boundaries. For each row, from left to right: raw histology patch, manual annotation, generated probability map and the output segmentation map. In both manual annotation and output segmentation map, different glands are assigned with different colors.

level, the shapes of different glands are well preserved, which is quite important for further diagnosis. In dealing with close glands, due to the constructed weight matrix, the deep model can effectively separate close glands although the gaps between close glands are slightly enlarged.

Several challenging cases are shown in Fig. 5. As we adopt a new shape similarity measurement instead of the pixel-wise cross-entropy measurement, the deep model should be able to detect boundaries more effectively. However, as we directly adopt the Otsu thresholding method to generate final segmentation maps for evaluation, some background pixels may not be successfully identified although the probability values are relatively lower than those neighboring gland pixels. One possible way to solve the problem is to search for an optimal threshold to generate the final segmentation map or to utilize an adaptive local thresholding scheme. These solutions will be explored in our future work as post-processing for performance improvement. Another issue is caused by the curve fitting method in the shape similarity measurement as shown in the second row of Fig. 5. Since the lengths of l_i and s_i in (1) are not required to be exactly the same, it is possible that although most boundary pixels are correctly identified, two glands may be wrongly connected by some discrete pixels. One solution to address the problem is to reduce the default length in the boundary segmentation process as discussed in Sect. 2.1.

Exemplar failure cases are shown in Fig. 6. The main reason for these failures is the shortage of corresponding training samples. One interesting observation is that although the deep model fails to identify some glands, the boundaries of

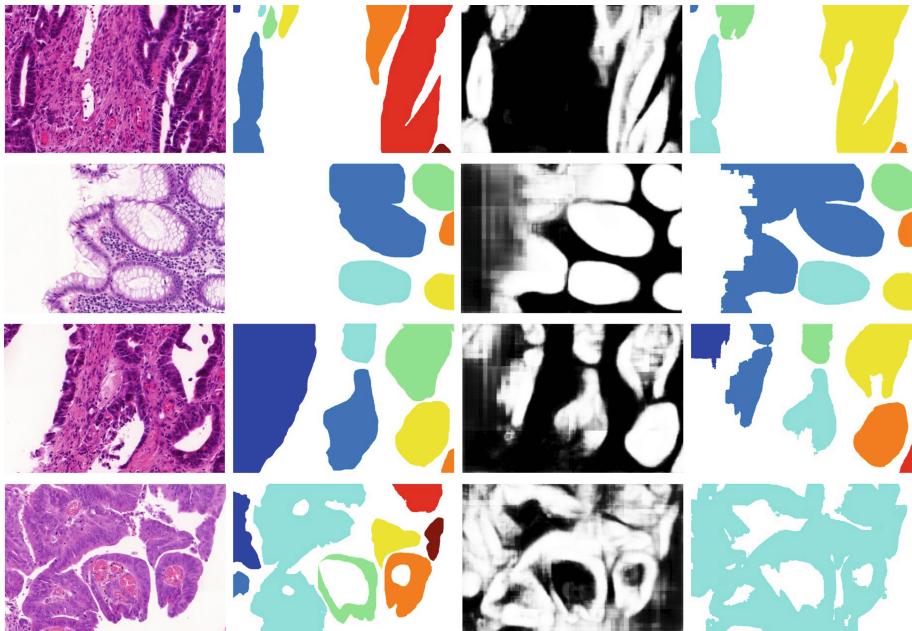


Fig. 6. Exemplar bad results generated by the proposed deep model. For each row, from left to right: raw histology patch, manual annotation, generated probability map and the output segmentation map. In both manual annotation and output segmentation map, different glands are assigned with different colors.

most segmented glands are quite similar to those annotated boundaries, which demonstrates the effectiveness of the proposed loss in shape preservation.

4 Conclusion

In this paper, we propose a deep model with a new shape-preserving loss that achieves the best overall performance on gland instance segmentation. Compared to the current best-known multi-channel/-task deep models, the shape-preserving loss enables one single deep model to generate accurate gland instance segmentation, which could largely reduce the model complexity and could be utilized in any other deep learning models for medical image segmentation.

References

1. Fleming, M., et al.: Colorectal carcinoma: pathologic aspects. *J. Gastrointest. Oncol.* **3**(3), 153–173 (2012)
2. Xu, Y., et al.: Gland instance segmentation using deep multichannel neural networks. *IEEE Trans. Med. Imaging* **64**(12), 2901–2912 (2017)

3. Xu, Y., et al.: Gland instance segmentation by deep multichannel neural networks. arXiv preprint [arXiv:1607.04889](https://arxiv.org/abs/1607.04889) (2016)
4. Chen, H., et al.: DCAN: deep contour-aware networks for accurate gland segmentation. In: CVPR, pp. 2487–2496 (2016)
5. Yan, Z., et al.: A skeletal similarity metric for quality evaluation of retinal vessel segmentation. IEEE Trans. Med. Imaging **37**(4), 1045–1057 (2017)
6. 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
7. Xie, S., Tu, Z.: Holistically-nested edge detection. In: ICCV, pp. 1395–1403 (2015)
8. Liu, Y., et al.: Richer convolutional features for edge detection. In: CVPR, pp. 5872–5881 (2017)
9. Sirinukunwattana, K.: Gland segmentation in colon histology images: the GlaS challenge contest. Med. Image Anal. **35**, 489–502 (2017)