HKUST SPD - INSTITUTIONAL REPOSITORY

Title Learning Hierarchical Semantic Correspondence and Gland Instance Segmentation

Authors Wang, Pei; Chung, Chi Shing

Source Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), v. 12436 LNCS, 2020, p. 603-613

Version Accepted Version

DOI 10.1007/978-3-030-59861-7_61

Publisher Springer Nature

Copyright © Springer Nature Switzerland AG 2020. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-030-59861-7_61

This version is available at HKUST SPD - Institutional Repository (<https://repository.ust.hk/ir>)

If it is the author's pre-published version, changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published version.

Learning Hierarchical Semantic Correspondence and Gland Instance Segmentation

Pei Wang^(✉) and Albert C. S. Chung

Lo Kwee -Seong Medical Image Analysis Laboratory, Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Kowloon, Hong Kong SAR
`pei.wang@connect.ust.hk`

Abstract. The morphology statistics of colon glands is a key feature for pathologists to diagnose colorectal cancer. Current gland instance segmentation methods show good overall performances, but accurate segmentation of extremely deformed glands in highly malignant cases or some rare benign cases remains to be challenging. In this paper, we propose a hybrid model that learns hierarchical semantic feature matching from histological pairs in an attentive process, where both spatial details and morphological appearances can be well preserved and balanced, especially for the glands with severe deformation. A consistency loss function is also introduced to enforce simultaneous satisfaction of semantic correspondence and gland instance segmentation on the pixel-level. The novel proposed model is validated on two publicly available colon gland datasets GlaS and CRAG. The model successfully boosts the segmentation performances on greatly mutated or deformed cases, and outperforms the state-of-the-art approaches.

Keywords: Hierarchical semantic correspondence · Consistency loss · Attentive process · Gland instance segmentation

1 Introduction

Colorectal cancer is the third most common cancer worldwide [16]. In clinical practice, the morphological statistics of intestinal glands is one of the primary features to colorectal cancer diagnosis and further treatment plan [3]. Automated gland segmentation is essential to provide morphological statistics to raise assessing efficiency and reliability, reduce inter- and intra-observer variability, and handle the ever-increasing image quantity.

Recently, CNN plays a significant role in the gland instance segmentation of histology images. To separate gland clusters, DCAN [4] performs gland segmentation and boundary detection in a unified network, and won the 2015 MICCAI GlaS challenge [15]. Based on the architecture of DCAN, MILD-Net [6] adopts maximal information loss to retain more spatial details in the feature extraction. A multi-channel multi-task network is proposed by Xu et al. [19, 20], where

Fig. 1. Examples of gland histology images. Great morphological changes in some benign cases and highly malignant cases remain to be a challenge in gland segmentation.

the foreground segmentation, edge detection, and object detection for glandular objects are conducted simultaneously. Besides, a loss function for boundary regularization is introduced by Yan et al. [21] to focus on the gland shape. Similarly, Qu et al. [13] abandoned pooling layers to improve the localization accuracy and applied a variance constraint on the loss function to explore more spatial features. DSE [17] is a nested network that first segments the gland and then refines the results by learning the inconsistency between ground truth and predicted masks. Furthermore, some methods also achieve competitive results by utilizing less manual annotations [22, 24, 25] and network quantization [18].

However, the accurate segmentation of gland instance with great diversity in shape, structure, and appearance remains to be a challenge, as shown in Fig. 1. Basically, current methods focus more on exploiting the boundaries [4, 19–21] and spatial information [6, 13], where the huge morphological changes can undermine their performances. Another important reason is that these greatly deformed cases take up a small proportion [5] of the population/dataset (a few benign cases and some highly malignant cases), which results in insufficient training samples for these complex cases. To address these problems, we propose a model that learns semantically related features from histological pairs with a hierarchical attention process, where the multi-level features from spatial information to morphological appearance can be emphasized and balanced even for complicated cases with severe gland deformation. A consistency loss is also proposed to enforce pixel-wise satisfaction of semantic feature correspondence and gland instance segmentation. The proposed method is validated on two publicly available colon gland datasets and outperforms the state-of-the-art approaches.

2 Methodology

A pair of histologic image pairs are fed into our model to learn the hierarchical semantic correspondence and gland instance segmentation simultaneously, as shown in Fig.2.

2.1 Feature Extraction and Matching

The source and target input images (I^s, I^t) are fed into a siamese network branch for extracting semantic features with shared learning weights. The feature extractor and decoder consist of dilated [23] residual blocks [7] following the U-Net

Fig. 2. Overview of the proposed model. A pair of source and target images are fed into the model, and features are extracted by a siamese network encoder. Pairwise feature matching and attentive modules are established hierarchically for instance segmentation. At training time, corresponding foreground masks are used to compute gland segmentation, semantic correspondence, and consistency loss terms.

architecture [14], and the feature maps of different layers contain information of different levels. For a certain layer of the feature extractor, the output features maps (f^s, f^t) with size $h \times w \times d$ for each image, can be represented as $h \times w$ grids of local features $f_{i,j}^s$ and $f_{u,v}^t$ in d -dimension. L2 normalization is applied to the individual d -dimensional features. The matching between local features of two images can be computed by the dot product operation. The matching scores are calculated by a normalized correlation function and defined as:

$$\mathcal{M} : \mathbb{R}^{h \times w \times d} \times \mathbb{R}^{h \times w \times d} \rightarrow \mathbb{R}^{h \times w \times h \times w}, \quad (1)$$

$$m_{ijuv} = \mathcal{M}(f^s, f^t)_{ijuv} = \frac{\langle f_{i,j}^s, f_{u,v}^t \rangle}{\sqrt{\sum_{i',j'} \langle f_{i',j'}^s, f_{u,v}^t \rangle^2}}. \quad (2)$$

The normalization operation emphasizes the similarity scores of highly-related local features and also penalizes the other ambiguous matches. The resulting $\mathbf{m} = \{m_{ijuv}\}$ contains all the normalized match scores between two features.

The soft argmax function [8, 10] is adopted to calculate the best matches based on the similarity scores. The differentiable soft argmax is applied to the 2-dimensional correlation map $\mathbf{m}$ for feature $f_{i,j}^s$ with coordinate (i, j) . To be more specific, the coordinates of semantic correspondence to feature $f_{i,j}^s$ are calculated by a weighted average of matching scores of all spatial locations (u, v) in the target feature maps. The matching probability $\mathbf{w}$ is assigned to different locations by applying softmax function to our correlation map $\mathbf{m}$. The matching function and corresponding probability map is defined as:

$$\mathcal{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \mathcal{T}(i, j) = \sum_{u,v} w_{i,j}(u, v) (u, v), \quad (3)$$

$$w_{i,j}(u, v) = \frac{\exp(m_{ijuv}/\tau)}{\sum_{u',v'} \exp(m_{ijuv'}/\tau)}. \quad (4)$$

τ is a temperature parameter. For higher temperature, all coordinates tend to have nearly the same probability, and the lower the temperature, the probability of the expected location is close to one.

2.2 Hierarchical Semantic Attention

After semantic correspondence scoring and matching, the feature maps of the target image from different layers are refined by an attention process. The attention weights for feature maps are defined by the matching probability of the correspondence between the source and target images. Given feature maps $f_{u,v}^t$ of the target image and the correspondence mapping $T(i, j) = (u, v)$, let $\alpha = \{\alpha_{i,j}\}$ be the attention probability distribution over the feature maps, we compute the attended feature $f_{u,v}^{att}$ by element-wise multiplication:

$$f_{u,v}^{att} = \alpha_{i,j} f_{u,v}^t, \quad \alpha_{i,j} = \frac{\exp(m_{ij}T(i,j)/\tau)}{\sum_{i',j'} \exp(m_{i'j'}T(i',j')/\tau)}. \quad (5)$$

The attention weight is formed by the similarity scores of semantic correspondence between feature maps and then gone through a softmax function. In the feature extractor, deeper layers have larger receptive fields and result in more features related to shape and appearance in the corresponding feature maps and attention maps. A series of hierarchical attention maps are generated to emphasize features from spatial to appearance level. In Fig. 3, we visualize the semantic correspondence of the input image pair with different feature levels and the corresponding segmentation performance.

2.3 Loss

The binary foreground masks of histology images are employed for the supervision of both segmentation and correspondence tasks in training the proposed network. The loss function can be expressed as:

$$\mathcal{L} = \lambda_{seg}\mathcal{L}_{seg} + \lambda_{cor}\mathcal{L}_{cor} + \lambda_{con}\mathcal{L}_{con}. \quad (6)$$

It consists of loss function for gland segmentation $\mathcal{L}_{seg}$, semantic correspondence $\mathcal{L}_{cor}$, and our proposed $\mathcal{L}_{con}$ for the consistency of these two tasks. We use cross entropy for the gland instance segmentation:

$$\mathcal{L}_{seg} = \mathbf{w} \mathcal{L}'_{seg} = - \sum_{u,v} w(u,v) \log p((u,v), y(u,v); \theta_{seg}), \quad (7)$$

$$w(u,v) = \max(w_0(d_0 - \text{dist}^2(u,v)), 1). \quad (8)$$

$p((u,v), y(u,v); \theta_{seg})$ represents the pixel-based softmax classification for true labels $y(u,v)$ of pixel (u,v) . A weight map $\mathbf{w}$ is applied to better identify the closing glands, where $\text{dist}(u,v)$ is the Euclidean distance of a pixel to the nearest gland. In our experiments w_0 and d_0 are set to 0.037 and 55 respectively.

Fig. 3. Visualization of hierarchical semantic correspondence on histological pairs (feature matchings are visualized with thresholding for a clearer view).

We adapt the mask consistence loss [11] to regularize the semantic matching on foreground only. The flow field $\mathcal{F}^s$ from source to target image can be expressed as $\mathcal{F}^s(i, j) = \mathcal{T}(i, j) - (i, j)$. Similarly, $\mathcal{F}^t$ is defined as $\mathcal{T}(u, v) - (u, v)$. Let M^s and M^t be the binary masks of source and target images. We estimate the source mask by warping [9] using the flow field as $\hat{M}^s = \mathcal{W}(M^t; \mathcal{F}^s)$. The warping operator can be formulated as $\mathcal{W}(M^t; \mathcal{F}^s)(i, j) = M^t((i, j) + \mathcal{F}^s(i, j))$. Finally, our loss for semantic correspondence can be expressed as:

$$\mathcal{L}_{cor} = \frac{1}{|N^s|} \sum_{i,j} (M^s(i, j) - \hat{M}^s(i, j))^2 + \frac{1}{|N^t|} \sum_{u,v} (M^t(u, v) - \hat{M}^t(u, v))^2, \quad (9)$$

where $|N|$ is the number of pixels in the mask. The function encourages matches between foreground pixels and penalizes other matching cases. For the semantic feature matching of different layers, we resize the masks and take the average of loss terms from different feature maps. Besides, due to the nature of histology images that contain multiple gland objects and complex background, the proposed model learns feature matching instead of warping the entire image.

Besides, we introduce a consistency loss to encourage simultaneous sanctification of the two tasks based on the target flow field and the softmax classification for segmentation, and with the same weight focusing on the boundary area:

$$\mathcal{L}_{con} = - \sum_{u,v} w(u, v) [(M^t(u, v) - \hat{M}^t(u, v))^2 + k] \log p((u, v), y(u, v); \theta_{seg}), \quad (10)$$

where k is a parameter for the consistency level of two tasks. In our cases we set k to 0.5. The function enforces correct prediction of foreground objects and precise matching within the same class for each pixel of the target images.

3 Experiments

We evaluate the proposed method on two public datasets for gland instance segmentation. The experiments demonstrate that our attentive model with hierarchical semantic matching and consistency loss substantially improves the performance of extremely complex cases and achieves state-of-the-art results.

Table 1. Performance on GlaS dataset in comparison with other methods.

Method	F1 Score		Object Dice		Object Hausdorff	
	Part A	Part B	Part A	Part B	Part A	Part B
Ours	0.931	0.872	0.929	0.882	27.143	74.175
DSE [17]	0.926	0.862	0.927	0.871	31.209	80.509
Quantization [18]	0.930	0.862	0.914	0.859	41.783	97.390
FullNet & varCE [13]	0.924	0.853	0.914	0.856	37.28	88.75
Mild-Net [6]	0.914	0.844	0.913	0.836	41.540	105.890
Shape Loss [21]	0.924	0.844	0.902	0.840	49.881	106.075
Suggestive Annotation [22]	0.921	0.855	0.904	0.858	44.736	96.976
Multi-channel [20]	0.893	0.843	0.908	0.833	44.129	116.821
DCAN [4]	0.912	0.716	0.897	0.781	45.418	160.347

Datasets & Evaluation. We validate the proposed model on GlaS dataset [15] from 2015 MICCAI challenge, and CRAG dataset [6] from Awan et al. [2]. GlaS consists of 85 training (37 benign (BN) and 48 malignant (MT)) and 80 test images (33 BN and 27 MT in Part A, 4 BN and 16 MT in Part B as on-site test of the challenge). In CRAG dataset there are 173 training and 40 testing images. The official metrics of GlaS are for different aspects: F1 score for object detection, *object-level* Dice index for instance segmentation, and *object-level* Hausdorff distance for glandular shape similarity [15]. The dense semantic correspondence is evaluated by the number of correctly labeled pixels: label transfer accuracy (LT-ACC) and the intersection-over-union (IoU) [11].

Training & Implementation Details. In the testing phase, we find the best matched image pairs by a simple and fast entropy-histogram approach [1]. In order to alleviate the over-fitting issue, we use random pairs from groups in training. Images with higher similarity are assigned into one group, and random image pairs are formed from the group for training. In our experiments, the images are augmented by elastic transformation, rotation, flip, Gaussian blur, and color distortion. Eventually, we randomly crop patches of size 480×480 . Our method is implemented with PyTorch [12]. We adopt Gaussian distribution ($\mu = 0, \sigma = 0.01$) for weight initialization and train with Adam optimization of initial learning rate 8×10^{-4} , with batch size of 2. For loss function, we assign the coefficients for segmentation, correspondence, and consistency as 1.5, 1.0 and 0.85. For both datasets, we split 20% of the training images for validation. The model is trained and tested on an NVIDIA Titan X Pascal for both datasets.

3.1 Results

Table 1 shows the comparative results on GlaS dataset. It includes methods using nested network DSE [17], and multi-channel multi-task model [20]; approaches

Fig. 4. Segmentation examples of benign cases on GlaS dataset (top three from part A, bottom two from part B).

Table 2. Ablation study on GlaS dataset.

Method	F1 Score		Object Dice		Object Hausdorff		LT-ACC		IoU	
	Part A	Part B	Part A	Part B	Part A	Part B	Part A	Part B	Part A	Part B
Baseline	0.923	0.854	0.916	0.862	40.456	88.463	-	-	-	-
+ Semantic Attn	0.927	0.860	0.923	0.870	34.421	79.245	0.948	0.863	0.876	0.808
+ Semantic Attn. & $\mathcal{L}_{con}$	0.931	0.872	0.929	0.882	27.143	74.175	0.962	0.908	0.894	0.831

with boundary segmentation like Shape Loss [21] and the challenge winner DCAN [4]; methods for more spatial information like Mild-Net [6] and FullNet with varCE [13]; and methods with suggestive annotation [22] and quantization [18]. As shown in Table 1, the proposed model outperforms all the other methods that are directly comparable. All three evaluation metrics for segmentation are promoted simultaneously, and substantial improvement lies in Part B for F1 score and object Dice, and object Hausdorff for Part A and B. These advancement indicates better performance for rather complex cases (75% of Part B are malignant cases) and better recognition of overall gland structures.

Figure 4 and Fig. 5 visualize the exemplar segmentation results of benign and malignant cases in GlaS dataset. The common issues in segmenting benign

Fig. 5. Segmentation examples of malignant cases on GlaS dataset (top two from part A, bottom three from part B).

Table 3. Performance on CRAG dataset in comparison with other methods.

Method	F1 Score	Object Dice	Object Hausdorff
Ours	0.840	0.898	109.02
DSE [17]	0.835	0.889	120.127
Mild-Net [6]	0.825	0.875	160.14
DCAN [4]	0.736	0.794	218.76

cases include the touching glands, the misleading background, and the structural changes in a few cases. As shown in Fig. 4, the proposed method shows better performances in handling the details and the recognition of overall glandular structures. Noticed that the baseline method is the proposed one without semantic feature matching and $\mathcal{L}_{con}$. For malignant cases in Fig. 5, the gland morphology appears to be more complicated, where the majority of them have no regular shape or structure. With the visual aids from the semantically related images, the morphological change is well captured by the proposed model.

Table 2 demonstrates the ablation study on GlaS dataset. The baseline method achieves promising performance. With the addition of the proposed

hierarchical semantic attention module, an 5% improvement can be observed on average for all three metrics of segmentation, and especially the Hausdorff distance that measures glandular shape similarity. With the consistency loss, the proposed method further promotes the results by 4% on average for segmentation and semantic correspondence simultaneously with better shape recognition. The hard cases are greatly improved in terms of the glandular shape.

Table 3 lists the comparative results on CRAG dataset. It is a more recent dataset and we have reported all results to the best of our knowledge. Similarly, the advancements in segmentation metrics and especially the Hausdorff distance validate the effectiveness of our proposed method in shape recognition.

In our experiments, it takes about 6 h to train the model, and 0.4s to 0.9s to segment the test image, depending on the image size. We have also tested the performance of the model with two decoders and performed simultaneous segmentation of both source and target images, which is more efficient but downgrades the overall performances. The possible reason could be the siamese encoder fails to extract sufficient features for both images for pixel-level classification, especially when the similarity of the histologic pairs are not high enough. This problem can be further discussed and studied in the future.

4 Conclusion

In this paper, we tackle the problem of morphology diversity in highly deformed gland cases by an attentive model with hierarchical semantic feature matching, where both spatial and appearance features are well preserved for accurate gland segmentation. A consistency loss is introduced to supervise the simultaneous fulfillment of feature matching and segmentation on pixel-level. The proposed model is validated on two public gland datasets and outperforms the state-of-the-art approaches with a major improvement in the greatly deformed cases.

References

1. Aljanabi, M.A., Hussain, Z.M., Lu, S.F.: An entropy-histogram approach for image similarity and face recognition. *Mathematical Problems in Engineering* **2018**, (2018)
2. Awan, R., Sirinukunwattana, K., Epstein, D., Jefferyes, S., Qidwai, U., Aftab, Z., Mujeeb, I., Snead, D., Rajpoot, N.: Glandular morphometrics for objective grading of colorectal adenocarcinoma histology images. *Sci. Reports* **7**(1), 16852 (2017)
3. Bosman, F.T., Carneiro, F., Hruban, R.H., Theise, N.D., et al.: WHO classification of tumours of the digestive system. No. 4 En., World Health Organization, (2010)
4. Chen, H., Qi, X., Yu, L., Heng, P.A.: DCAN: deep contour-aware networks for accurate gland segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2487–2496 (2016)
5. Fleming, M., Ravula, S., Tatishchev, S.F., Wang, H.L.: Colorectal carcinoma: pathologic aspects. *J. Gastrointestinal Oncol.* **3**(3), 153 (2012)
6. Graham, S., et al.: Mild-net: minimal information loss dilated network for gland instance segmentation in colon histology images. *Med. Image Anal.* **52**, 199–211 (2019)

7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
8. Honari, S., Molchanov, P., Tyree, S., Vincent, P., Pal, C., Kautz, J.: Improving landmark localization with semi-supervised learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1546–1555 (2018)
9. Jaderberg, M., Simonyan, K., Zisserman, A., et al.: Spatial transformer networks. In: Advances in Neural Information Processing Systems, pp. 2017–2025 (2015)
10. Kendall, A., et al.: End-to-end learning of geometry and context for deep stereo regression. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 66–75 (2017)
11. Lee, J., Kim, D., Ponce, J., Ham, B.: Sfnets: learning object-aware semantic correspondence. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2278–2287 (2019)
12. Paszke, A., et al.: Pytorch: an imperative style, high-performance deep learning library. In: Advances in Neural Information Processing Systems, pp. 8024–8035 (2019)
13. Qu, H., Yan, Z., Riedlinger, G.M., De, S., Metaxas, D.N.: Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11764, pp. 378–386. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32239-7_42
14. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
15. Sirinukunwattana, K., Pluim, J.P., Chen, H., Qi, X., Heng, P.A., Guo, Y.B., Wang, L.Y., Matuszewski, B.J., Bruni, E., Sanchez, U., et al.: Gland segmentation in colon histology images: the glas challenge contest. *Med. Image Anal.* **35**, 489–502 (2017)
16. Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., Jemal, A.: Global cancer statistics, 2012. *CA: A Cancer J. Clin.* **65**(2), 87–108 (2015)
17. Xie, Y., Lu, H., Zhang, J., Shen, C., Xia, Y.: Deep segmentation-emendation model for gland instance segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11764, pp. 469–477. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32239-7_52
18. Xu, X., et al.: Quantization of fully convolutional networks for accurate biomedical image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8300–8308. IEEE (2018)
19. Xu, Y., Li, Y., Liu, M., Wang, Y., Lai, M., Eric, I., Chang, C.: Gland instance segmentation by deep multichannel side supervision. In: Ourselin, S., Joskowicz, L., Sabuncu, M., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 496–504. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_57
20. Xu, Y., Li, Y., Wang, Y., Liu, M., Fan, Y., Lai, M., Eric, I., Chang, C.: Gland instance segmentation using deep multichannel neural networks. *IEEE Trans. Biomed. Eng.* 64(12), 2901–2912 (2017)

21. Yan, Z., Yang, X., Cheng, K.T.T.: A deep model with shape-preserving loss for gland instance segmentation. In: Frangi, A., Schnabel, J., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11071, pp. 138–146. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00934-2_16
22. Yang, L., Zhang, Y., Chen, J., Zhang, S., Chen, D.Z.: Suggestive annotation: a deep active learning framework for biomedical image segmentation. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 399–407. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_46
23. Yu, F., Koltun, V.: Multi-scale context aggregation by dilated convolutions. arXiv preprint, (2015) [arXiv:1511.07122](https://arxiv.org/abs/1511.07122)
24. Zhang, Y., Yang, L., Chen, J., Fredericksen, M., Hughes, D.P., Chen, D.Z.: Deep adversarial networks for biomedical image segmentation utilizing unannotated images. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 408–416. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_47
25. Zheng, H., et al.: Biomedical image segmentation via representative annotation. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 5901–5908 (2019)