MULTI-DIRECTIONAL DISTANCE REPRESENTATION WITH BOUNDARY REFINEMENT FOR ROBUST CLUSTERED NUCLEI SEGMENTATION: A SOLUTION TO CONIC2022 CHALLENGE

Shuolin Liu

BeiJing, China

Nuclei instance segmentation within histology images enables the extraction of interpretable cell-based features that can be used in downstream explainable models in computational pathology (CPath). However, this task faced with two major challenges. One lies in the overlapping/clustered nuclei, while the other one lies in the varying size and shape of different structures in histology images. To address these challenges, we propose a distance-based instance segmentation model with boundary refinement unit for nuclei segmentation. Firstly, distance images are utilised to help model to recognize the boundary of each neclei, which are constructed from instance masks and records the distances of neclei pixels to their centres of mass in different directions. Secondly, we further employ a boundary refinement unit to refine each segmented instance, where different size of nuclei are processed under the same scale. Five-fold cross validation on CoNIC2022 challenge dataset is used to evaluate the method. Overall,the binary panoptic quality (PQ) is 0.69 and the multi-class panoptic quality (mPQ+) of six types of nuclei is 0.58. Experimental results demonstrate propose method can achieve robust nuclei instance segmentation within histology images, particularly in areas with clustered instances.

1. INTRODUCTION

Nuclei segmentation is a crucial task in computational pathology (CPath), as it provides rich spatial and morphometric information regarding nuclei. However, automatic nuclei segmentation remains challenging. Firstly, nuclei are usually clustered together. Secondly, the size, shape of different nucleus varying greatly.

The key to seperate clustered nuclei is detecting the boundary of each nucleus. One popular scheme is to segment the contours first and then subtract contours from predicted foreground to obtain final segmentation. In order to segment contour accurately, chen et al [1] adopted two independent decoders for U-Net [7], one for foreground segmentation and another for contour segmentation. Since there is no interaction between these two decoders, inconsistency may exist between segmented foreground and segmented contour. In order to make use of the correlation between two decoders, BES-Net[2] , CIA-Net [3], BRP-Net [4] further introduced uni-directional and bi-directional information transmission, respectively, which means one decoder obtains extra features from the other one. Another popular scheme is to employ distance modeling with morphological post-processing to separate clustered nuclei. Hover-Net [5] employ horizontal and vertical distance representation for cluster object separation, which achieved state-of-the-art results on multi-type nuclei segmentation and classification tasks.

In this paper, we propose a novel solution for nuclei segmentation, which is an effective extension of Hover-Net [5]. This solution comprises two stages: one stage to obtain instance proposals and another stage to refine the segmented boundary. In the first stage, we implement the multi-directional distance images to separate clustered nuclei. In the second stage, a boundary refinement unit is employed to obtain more accurate boundary for each segmented instance. Besides, a generic U-Net [7] with densenet-121 [6] as backbone is employed to further classify each segmented instance. We conducted experiments on CoNIC 2022 challenge dataset [8, 9], from which we can conclude that proposed approach significantly improve the nuclei segmentation accuracy.

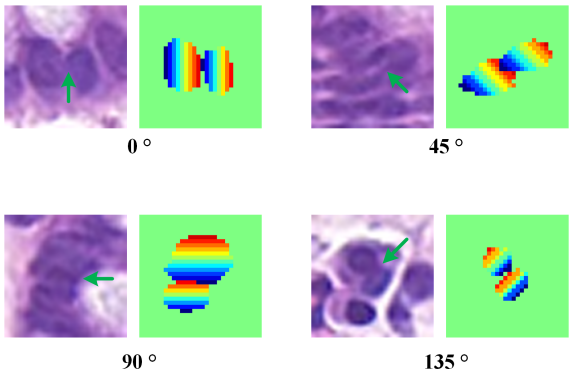

Fig. 1. Examples of clustered nuclei and their distance representations. Distance images with directions of { 0°, 45°, 90°, 135°} are displayed in the figure, where pixel values outside region of interest are set to 0 for better visualization.

Correspondence to: Shuolin Liu(1980073622@qq.com).

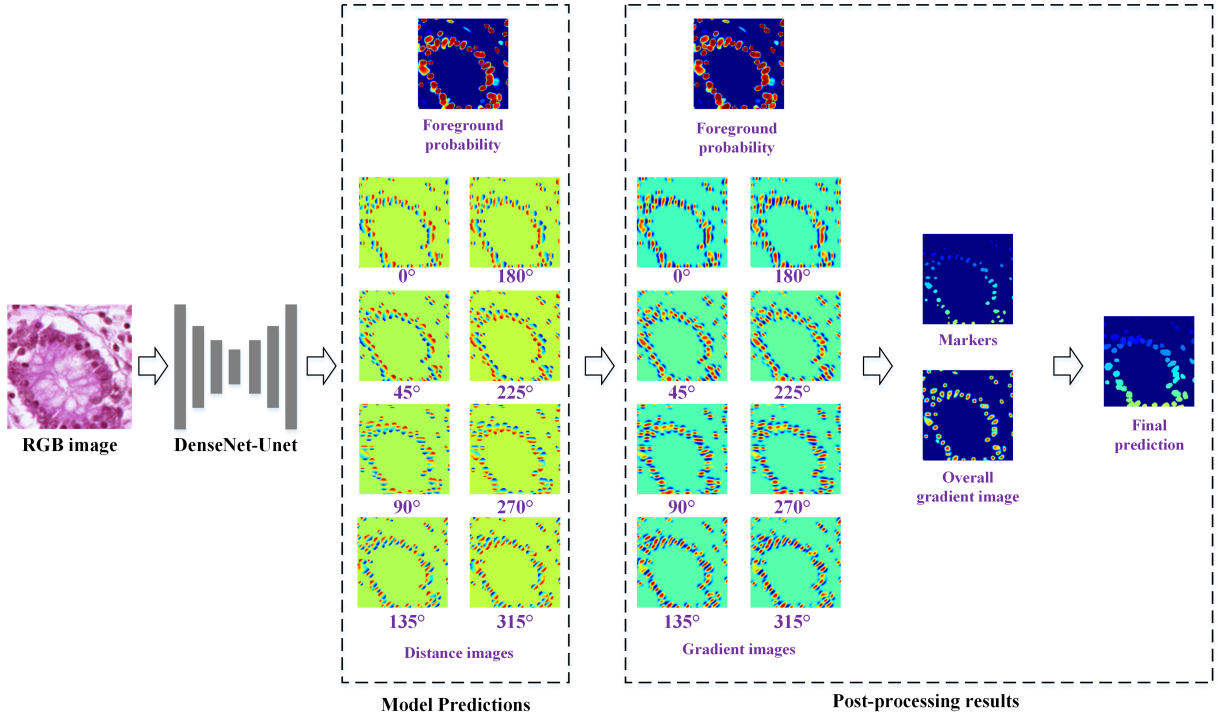

Fig. 2. Schematic diagram of the proposed framework for instance proposal generation.

2. METHODS

Similar to BRP-Net [4], our method for nuclei instance segmentation consists of two parts, one for obtaining instance proposals and another for instance-wise boundary refinement.

2.1. Instance Proposal Generation

Our scheme for obtaining high quality instance proposals is to predict distance information of nuclear pixels and then employ a tailored post-processing pipeline to segment nuclear instances. Extant approach Hover-Net[5] have integrated horizontal and vertical distance images into instance segmentation, however, as shown in Fig1, as nuclei may cluster in different directions, this approach may not have sufficient capacity to recognize the spatial relations between clustered instances. Accordingly, we extend Hover-Net' distance images from 2 directions to 8 directions(0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°).The distance images record the distances of nuclei pixels to their centres of mass in different directions.

2.1.1. Model Architecture

The model architecture is a variant of U-Net [7] with DenseNet-121 [6] as backbone.The input to the model is the histology image in RGB format, and the model outputs are the predicted foreground and eight predicted distance images. Compared to the standard DenseNet implementation, we reduce the total down-sampling factor from 32 to 8 by using a stride of 1 in the first convolution and removing the subsequent pooling operation. The number of dense units within each stage is 6, 12, 24, and 16 that are applied at down-sampling levels 1, 2, 4 and 8 respectively. The up-sampling path is employed to gradually recover the resolution of feature images and generates a high-resolution prediction. The decoder block mainly consists of repeating up-sampling layers, skip concatenations, 3×3 convolution layers, batchnormalization(BN)[10] layers, and ReLUs. In order to obtain richer details, the encoder layers with $\{1, 1/2, 1/4\}$ of the original image size are passed to the decoder block by skip concatenations since earlier high-resolution feature images from encoder can help to refine the information of locations.We adopted bilinear scaling followed by 3×3 convolution layers to up-sample the features images. Based on the decoder module, feature images are gradually up-sampled to the original resolution, and the predicted distance images and predicted foreground are generated after 1×1 convolution layers. We set the number of channels in decoder to 64 based on experiments.

2.1.2. Post-processing pipeline

As show in Fig. 2, the post-processing pipeline is utilized to transform predicted distance images and foreground probability into instance masks. We follow the similar rules outlined in HoverNet. Firstly, pixels with foreground probabilities larger than σ will be considered to be foreground pixels.

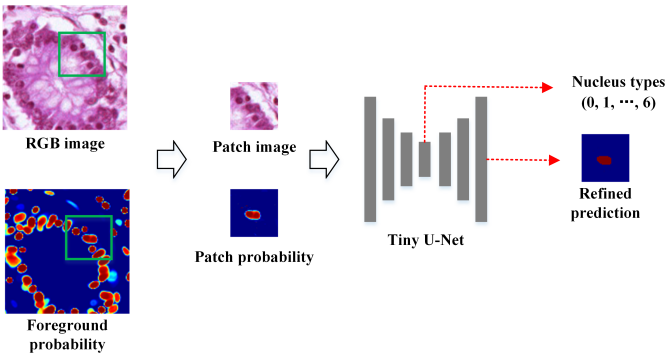

Fig. 3. Schematic diagram of the proposed boundary refinement unit.

We select a very small threshold σ=0.25 since boundary refinement unit in the next stage has some abilities to remove false positives generated from this stage. Then, 8 gradient images are firstly calculated from 8 distance images independently using different sobel kernels and then normalized to (0, 1) and merged into an overall gradient image. This operation could be very fast with GPU implementation. As we can see, pixels with high gradient values intend to be the center area pixels of nuclei. Therefore, we compute markers using a threshold function that acts on overall gradient image and sets values above 0.5 to 1 or 0 otherwise. With the markers, gradient image and predicted mask, the final predicted instance mask are generated after marker-controlled watershed algorithm.

2.2. Boundary refinement unit

As shown in Fig. 3, in order to handle the varying size of nuclei, the boundary refinement unit precesses segmented instances using the same sacle. We crop one square 48 $\times$ 48 patch containing each segmented instance. If an instance larger than 48, it will be resized to 48 $\times$ 48. Inputs to the model include the patch image, and the probability map that predicted by previous stage. To relieve the influence of background, elements in the probability map that fall outside of the dilated proposal are set to zero., we use a tiny 4-stage U-Net model for this unit. The number of model channels are 64, 128, 256, 512 respectively for each stage. To capture more contextual feature, a classification branch that classify nuclei into 7 types is employed for deep supervision. This branch simply consists of a single fully connected layer and will be removed during testing.

2.3. Nuclei classification

We simply use the same DenseNet121-U-Net for neclei classification. The input to the model is the histology image in RGB format, and the model outputs is 7-class predicted probability map. There is no special design for classification task.

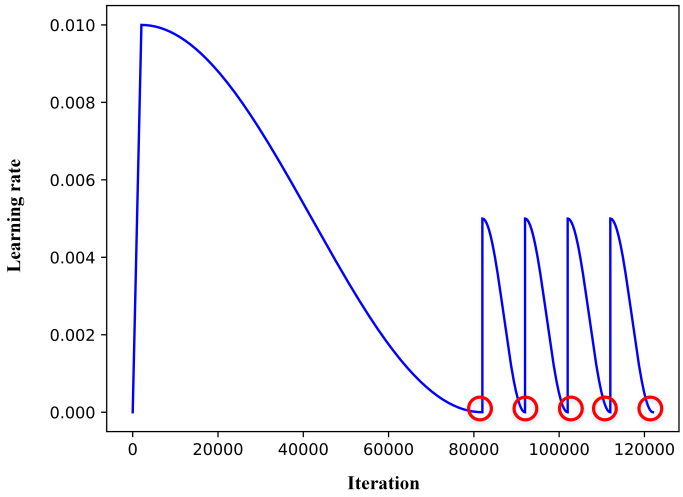

Fig. 4. Schematic diagram of the learning rate schedule.Red circles indicate where we employ the weight averaging.

3. EXPERIMENTS

We conduct five-fold cross validation on CoNIC 2020 chanllenge dataset [8, 9], which consists of 431,913 instances of six type nuclei. Models are trained using original histology images and validated using challenge patches in order to fairly compared with other methods. We strictly split the challenge patches at the patient level to make sure there is no data-leakage.

Binary panoptic quality(PQ) and multi-class PQ (mPQ+) described in [9] are employed as the evaluation metrics.Source code for evaluation could be found in [1].

3.1. Implementation Details

For instance proposals generation and classification, we employed similar loss functions as Hover-Net, which consists of CE loss, Dice loss and L1 loss. Weight of the L1 loss is set to 2. Models are trained using 128x128 randomly cropped images with batch size of 26 and tested using 256x256 images with batch size of 8. For boundary refinement, we use Dice loss and CE loss for optimization. Models are trained and tested using 48x48 patches with batch size of 128.

During training, data augmentation including flip, scaling, rotation, color jitter, motion blur, median blur, gaussian blur, gaussian noise was applied to all methods. During testing, flip and rotation are applied to instance proposal generation and classification model.

As shown in Fig. 4, SGD with 4-circle stochastic weights averaging (SWA) [11] were used to train all the models. The

[1] https://github.com/TissueImageAnalytics/CoNIC

Table 1. Ablation study results on a single fold. "BL" denotes the baseline, "MD" denotes the multi-directional distance image, "BR" denotes the boundary refinement unit, "SWA" denotes the stochastic weights averaging, "TTA" denotes the test-time-augmentation

Methods	Multi-PQ+							Binary PQ
	Neutrophil	Epithelial	Lymphocyte	Plasma	Eosinophil	Connective	Mean	
BL	0.319	0.621	0.725	0.578	0.395	0.637	0.546	0.663
BL+MD	0.321	0.631	0.730	0.582	0.398	0.641	0.551	0.673
BL+MD+BR	0.308	0.657	0.753	0.599	0.403	0.658	0.563	0.700
BL+MD+BR+SWA	0.337	0.663	0.756	0.604	0.420	0.663	0.574	0.704
BL+MD+BR+SWA+TTA	**0.351**	**0.665**	**0.759**	**0.605**	**0.417**	**0.670**	**0.578**	**0.706**

Table 2. Quantitative results of 5-fold cross validation

Methods	Multi-PQ+							Binary PQ
	Neutrophil	Epithelial	Lymphocyte	Plasma	Eosinophil	Connective	Mean	
Fold 0	0.351	0.665	0.759	0.605	0.417	0.670	0.578	0.706
Fold 1	0.472	0.662	0.736	0.602	0.466	0.652	0.598	0.688
Fold 2	0.315	0.646	0.770	0.579	0.418	0.662	0.565	0.678
Fold 3	0.331	0.673	0.733	0.594	0.437	0.668	0.573	0.700
Fold 4	0.366	0.661	0.733	0.595	0.489	0.673	0.586	0.695
Mean	0.367	0.661	0.746	0.594	0.445	0.664	0.580	0.693

basic learning rate is 0.01 and adapted following 2k warm-up and 120k cosine annealing scheduler. The optimization procedure is stopped at 122k iterations. All models are implemented using a single 3090 GPU with 24 GB memory.It takes around 25 hours to train the entire framework.

3.2. Ablation Study

In order to evaluate the effect of each component proposed in this paper, we conducted ablation studies on fold 0. The baseline is the instance proposal generation model that predicts foreground, horizontal and vertical distance images like Hover-Net. The results are exhibited in Tab. 1. As we can see that utilizing the multi-directional distance image and the boundary refinement unit improve the segmentation accuracy greatly. Further combining the SWA, test-time augmentation can slightly boost the performance.

3.3. Quantitative results of 5-fold cross validation

The provided results analysis is based on the 5-fold cross validation results on the training set. Tab. 2 illustrates the results of 5-fold cross validation. While high PQ and mPQ+ scores are obtained for epithelial, lymphocyte, plasma and connective, PQ and mPQ+ scores for neutrophil and eosinophil indicate unsatisfactory performance. As we can notice that the PQ is much better than mPQ+. As shown in Fig. 5,proposed method can obtain accurate instance segmentation results even on clustered area, which means the limitation of the framework may lies in the classification model.

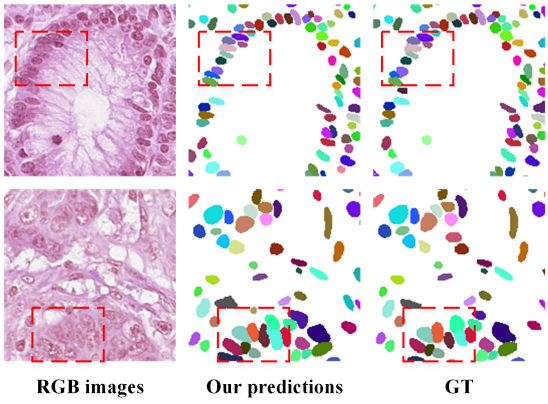

Fig. 5. Visualization results.From left to right are input RGB images, our predictions and ground truth(GT).

4. CONCLUSION

This paper presents a novel framework for nuclei segmentation with clustered instances. To accurately split clustered nucleis, we introduced multi-directional distance representation into nuclei proposal generation. To further boost the performance of nuclei segmentation, boundary refinement unit is designed to learn contextual features in the same scale. Experimental results demonstrate the proposed method can work well on nuclei segmentation and classfication. As such, it shows highly potential clinical values in CPATH.

5. REFERENCES

[1] Hao Chen, Xiaojuan Qi, Lequan Yu, and Pheng-Ann Heng, "DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 2016, pp. 2487–2496, IEEE.

[2] Hirohisa Oda, Holger R. Roth, Kosuke Chiba, Jure Sokolić, Takayuki Kitasaka, Masahiro Oda, Akinari Hinoki, Hiroo Uchida, Julia A. Schnabel, and Kensaku Mori, "BESNet: Boundary-Enhanced Segmentation of Cells in Histopathological Images," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, Alejandro F. Frangi, Julia A. Schnabel, Christos Davatzikos, Carlos Alberola-López, and Gabor Fichtinger, Eds., vol. 11071, pp. 228–236. Springer International Publishing, Cham, 2018, Series Title: Lecture Notes in Computer Science.

[3] Yanning Zhou, Omer Fahri Onder, Qi Dou, Efstratios Tsougenis, Hao Chen, and Pheng-Ann Heng, "CIA-Net: Robust Nuclei Instance Segmentation with Contour-Aware Information Aggregation," in *Information Processing in Medical Imaging*, Albert C. S. Chung, James C. Gee, Paul A. Yushkevich, and Siqi Bao, Eds., vol. 11492, pp. 682–693. Springer International Publishing, Cham, 2019, Series Title: Lecture Notes in Computer Science.

[4] Shengcong Chen, Changxing Ding, and Dacheng Tao, "Boundary-Assisted Region Proposal Networks for Nucleus Segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Anne L. Martel, Purang Abolmaesumi, Danail Stoyanov, Diana Mateus, Maria A. Zuluaga, S. Kevin Zhou, Daniel Racoceanu, and Leo Joskowicz, Eds., vol. 12265, pp. 279–288. Springer International Publishing, Cham, 2020, Series Title: Lecture Notes in Computer Science.

[5] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot, "Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," *Medical Image Analysis*, vol. 58, pp. 101563, Dec. 2019.

[6] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, July 2017, pp. 2261–2269, IEEE.

[7] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, Eds., vol. 9351, pp. 234–241. Springer International Publishing, Cham, 2015, Series Title: Lecture Notes in Computer Science.

[8] Simon Graham, Mostafa Jahanifar, Ayesha Azam, Mohammed Nimir, Yee-Wah Tsang, Katherine Dodd, Emily Hero, Harvir Sahota, Atisha Tank, Ksenija Benes, Noorul Wahab, Fayyaz Minhas, Shan E Ahmed Raza, Hesham El Daly, Kishore Gopalakrishnan, David Snead, and Nasir Rajpoot, "Lizard: A Large-Scale Dataset for Colonic Nuclear Instance Segmentation and Classification," in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, BC, Canada, Oct. 2021, pp. 684–693, IEEE.

[9] Simon Graham, Mostafa Jahanifar, Quoc Dang Vu, Giorgos Hadjigeorghiou, Thomas Leech, David Snead, Shan E. Ahmed Raza, Fayyaz Minhas, and Nasir Rajpoot, "CoNIC: Colon Nuclei Identification and Counting Challenge 2022," *arXiv:2111.14485 [cs]*, Nov. 2021, arXiv: 2111.14485.

[10] Sergey Ioffe and Christian Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv:1502.03167 [cs]*, Mar. 2015, arXiv: 1502.03167.

[11] Pavel Izmailov, Dmitrii Podoprikhin, Timur Garipov, Dmitry Vetrov, and Andrew Gordon Wilson, "Averaging Weights Leads to Wider Optima and Better Generalization," *arXiv:1803.05407 [cs, stat]*, Feb. 2019, arXiv: 1803.05407.