Weakly Supervised Deep Nuclei Segmentation using Points Annotation in Histopathology Images

Hui Qu[1] HUI.QU@CS.RUTGERS.EDU
Pengxiang Wu[1] PW241@CS.RUTGERS.EDU
Qiaoying Huang[1] QH55@CS.RUTGERS.EDU
Jingru Yi[1] JY486@CS.RUTGERS.EDU
Gregory M. Riedlinger[2] GR338@CINJ.RUTGERS.EDU
Subhajyoti De[2] SD948@CINJ.RUTGERS.EDU
Dimitris N. Metaxas[1] DNM@CS.RUTGERS.EDU
[1] *Department of Computer Science, Rutgers University, Piscataway, NJ 08854, USA.*
[2] *Rutgers Cancer Institute, 195 Little Albany St, New Brunswick, NJ 08901, USA.*

Abstract

Nuclei segmentation is a fundamental task in histopathological image analysis. Typically, such segmentation tasks require significant effort to manually generate pixel-wise annotations for fully supervised training. To alleviate the manual effort, in this paper we propose a novel approach using points only annotation. Two types of coarse labels with complementary information are derived from the points annotation, and are then utilized to train a deep neural network. The fully-connected conditional random field loss is utilized to further refine the model without introducing extra computational complexity during inference. Experimental results on two nuclei segmentation datasets reveal that the proposed method is able to achieve competitive performance compared to the fully supervised counterpart and the state-of-the-art methods while requiring significantly less annotation effort. Our code is publicly available[1].

Keywords: Nuclei segmentation, Weak supervision, Deep learning, Voronoi diagram, Conditional random field.

1. Introduction

Nuclei segmentation is a critical step in the automatic analyses of histopathology images, because the nuclear features such as average size, density and nucleus-to-cytoplasm ratio are often related to the clinical diagnosis and management of cancer. Modern deep learning based nuclei segmentation methods (Xing et al., 2016; Kumar et al., 2017; Naylor et al., 2017, 2018; Mahmood et al., 2018; Janowczyk and Madabhushi, 2016; Qu et al., 2019) have achieved better performance than traditional approaches such as watershed segmentation (Veta et al., 2013) and graph-based segmentation (Al-Kofahi et al., 2010). However, the fully supervised training of deep neural networks in these methods requires a large amount of pixel-wise annotated data, which are difficult to collect because assigning a nucleus/background class label to every pixel in the image is time-consuming and requires specific domain knowledge. Therefore, methods using weak annotations are needed to reduce the annotation burden.

1. The code can be found at: https://github.com/huiqu18/WeaklySegPointAnno

There have been various methods using weak annotations in image segmentation. For natural images, weak annotations include image-level tags (Papandreou et al., 2015; Pathak et al., 2015), scribbles (Lin et al., 2016), points (Bearman et al., 2016) and bounding boxes (Dai et al., 2015; Khoreva et al., 2017; Rajchl et al., 2017). Image-level tags are the class information of objects, which are not used in medical image segmentation where object classes in images are usually fixed (e.g., nuclei and background in our task). Scribbles annotation, which requires at least one scribble for every object, is not suitable for our task due to the small size and large number of nuclei. The objectiveness prior in the points supervision work (Bearman et al., 2016) is not working here since nuclei are small and thus the prior is inaccurate. Bounding boxes are more well defined and are also commonly adopted in medical images (Yang et al., 2018; Zhao et al., 2018). However, it is still time-consuming and difficult to label an image using bounding boxes for hundreds of nuclei, especially when the density is high. Kervadec et al. (Kervadec et al., 2019) used a small fraction of full labels and imposed a size constraint in the loss function, which achieved good performance but is not applicable for multiple objects of a same class. Different from existing methods, in this work we propose to employ points annotation for nuclei segmentation. All a pathologist needs to do is mark the location of every nucleus with a point. Our method is efficient and more annotation-friendly, and to the best of our knowledge, this is the first time points annotation has been successfully applied to nuclei segmentation.

In practice, the points annotation itself is not sufficient to directly supervise the training of neural networks. To address this problem, we take advantage of the original image and the shape prior of nuclei to derive two types of coarse labels from the points annotation using the Voronoi diagram and the k-means clustering algorithm. The Voronoi diagram was ever used in nuclei detection (Kost et al., 2017) for training sample selection, but here we utilize it to generate the coarse labels for nuclei segmentation, which is a different and much harder task. These two types of coarse labels are then used to train a deep convolutional neural network (CNN) with the cross entropy loss.

A common problem in various weakly supervised segmentation tasks is that the key information near the object boundaries is missing. Therefore, post-processing like the dense conditional random field (CRF) (Chen et al., 2015) or graph search (Yang et al., 2018) is needed to refine the object boundaries, at the expense of increased processing time. Inspired by Tang et al.'s work (Tang et al., 2018), we utilize the dense CRF in the loss function to fine-tune the trained model rather than add a post-processing step, thereby leading to a more efficient model as the loss is no longer needed during inference. This property makes our method more preferable in nuclei segmentation of large Whole Slide Images.

In summary, the contributions of our work include:

- To the best of our knowledge, we are the first to successfully utilize the points annotation for nuclei segmentation in histopathology images.
- We present a new method for deriving two types of informative pixel-level labels from points label using the Voronoi diagram and k-means clustering algorithm, and employ the dense CRF loss for model refinement in nuclei segmentation.
- We show that our approach achieves competitive segmentation performance on two nuclei segmentation datasets. The accuracy is comparable to that obtained with full supervised approaches.

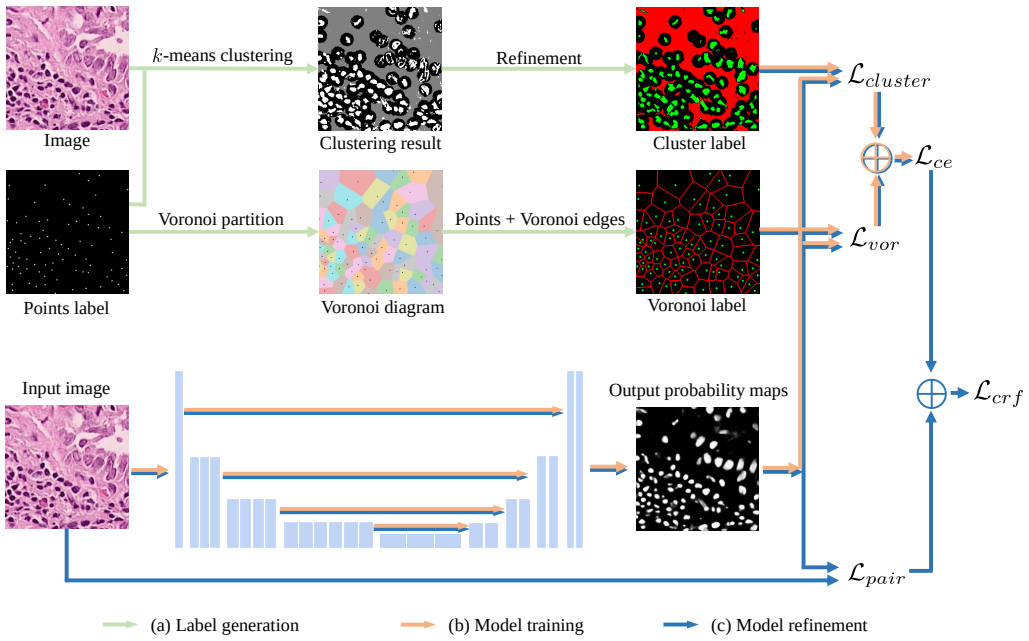


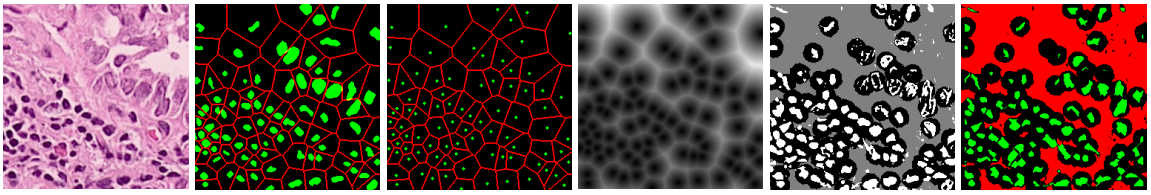
Figure 1: Overview of the proposed approach. (a) Label generation. The Voronoi label and cluster label are generated using the points label and original image. The green, red and black colors indicate nuclei, background and ignored pixels, respectively. (b) Model training using the cross entropy loss. (c) Model refinement using the CRF loss.

2. Methods

In this section we describe our approach in detail. In particular, our point-level supervision for training a nuclei segmentation model consists of three parts: (1) coarse pixel-level labels generation using points annotation; (2) segmentation network training with coarse labels; (3) model refinement using the dense CRF loss.

2.1. From point-level to pixel-level labels

The point-level labels cannot be used directly for the training of a CNN with the cross entropy loss due to the lack of (negative) background labels since all annotated points belong to the (positive) nuclei category. To solve this issue, the first step is to exploit the information we have to generate useful pixel-level labels for both classes. We have the following observations: (1) Each point is expected to be located or close to the center of a nucleus, and the shapes of most nuclei are nearly ellipses, i.e., they are convex. (2) The colors of nuclei pixels are often different from the surrounding background pixels. Based on these observations, we propose to utilize the Voronoi diagram and k-means clustering methods to produce two types of pixel-level labels.



(a) image (b) true mask (c) Voronoi label (d) distance map (e) cluster result (f) cluster label

Figure 2: Label generation. (a) original image, (b) ground-truth nuclei masks (in green) and Voronoi edges (in red), (c) Voronoi label, (d) distance map, (e) clustering result, (f) cluster label (green: nuclei, red: background, black: ignored).

2.1.1. VORONOI LABELS

Voronoi diagram is a partitioning of a plane into convex polygons (Voronoi cells) according to the distance to a set of points in the plane. There is exactly one point (seed point) in each cell and all points in a cell are closer to its seed point than other seed points. In our task, the annotated points of an image can be treated as seed points to calculate the Voronoi diagram, see Fig. 1. For each cell, assuming that the corresponding nucleus is located within the cell, then the Voronoi edges separate all nuclei well and the edge pixels belong to the background. This assumption holds for most of the nuclei because the points are around the centers and nuclear shapes are nearly convex (Fig. 2(b)).

Treating the Voronoi edges as background pixels and the annotated points (dilated with a disk kernel of radius 2) as nuclei pixels, we obtain the Voronoi point-edge label (Fig. 2(c)). All other pixels are ignored during training. Note that although the pixels on the Voronoi edge between two touching nuclei may not necessarily be background, the edges are still helpful in guiding the network to separate the nuclei. The Voronoi labels aim to segment the central parts of nuclei and are not able to extract the full masks, because they lack the information of nuclear boundaries and shapes. To overcome the weakness, we generate another kind of labels that contain this information as a complement.

2.1.2. CLUSTER LABELS

Considering the difference in colors between nuclei and background pixels, it is feasible to perform a rough segmentation using clustering methods. We choose the k-means clustering algorithm to extract both nuclei and background pixels from the original image, and produce the cluster labels based on the results. Given an image x with N pixels $(x_1, x_2, \cdots, x_N)$, k-means clustering aims to partition the N pixels into k clusters $S = (S_1, S_2, \cdots, S_k)$ according to the feature vector f_{x_i} of each pixel x_i, such that the sum of within-cluster variances is minimized :

$$\arg\min_{S} \sum_{i=1}^{k} \sum_{x \in S_i} \|f_x - c_i\|^2 . \tag{1}$$

We use k-means to divide all pixels into $k = 3$ clusters: nuclei, background and ignored. The cluster that has maximum overlap with points label is considered as nuclei, and the cluster that has minimum overlap with the dilated points label is considered as background. The remaining one is the ignored class. The pixels of ignored class are often located around the nuclear boundaries, which are hard for a clustering method to assign correct labels.

For the feature vector f, color is the straightforward choice. However, clustering with color will result in wrong assignments for pixels inside some nuclei that have non-uniform colors. To cope with this issue, we propose to add a distance value in the feature vector. In a distance map (Fig. 2(d)), each value indicates the distance of that pixel to the closest nuclear point and therefore incorporates the position information. In particular, the pixels that belong to nuclei should be close enough to points in the label while background pixels are expected to be relatively far from those points. The distance map can be calculated by the distance transform of the complement image of points label. Combining the distance value d_i with the RGB color values (r_i, g_i, b_i) as the feature vector $f_{x_i} = (\hat{d}_i, \hat{r}_i, \hat{g}_i, \hat{b}_i)$ and performing k-means clustering, we obtain the initial cluster labels (Fig. 2(e)). $\hat{d}_i$ is the clipped value by truncating large values to 20 and $\hat{r}_i$, $\hat{g}_i$, $\hat{b}_i$ are scaled color values such that each element in the feature vector has similar range. The final cluster label (Fig. 2(f)) is generated by refining the clustering result with morphological opening operation. The cluster labels have more shape information about the nuclei compared to Voronoi label, but may contain more errors and uncertainties. We argue that these two types of labels are complementary to each other and would jointly lead to better results.

2.2. Training deep neural networks with pixel-level labels

Once we have the pixel-level labels, we are able to train a deep convolutional neural network for nuclei segmentation. The network (shown in Fig. 1) we use is a modified version of U-net (Ronneberger et al., 2015). We replace the encoder part of U-net with the convolution layers of ResNet34 (He et al., 2016), which is more powerful in representation ability and can be initialized with pretrained parameters from image classification task on ImageNet (Russakovsky et al., 2015). The network outputs two probability maps of background and nuclei, which are used to calculate two cross entropy losses with respect to the cluster label $\mathscr{L}_{cluster}$ and Voronoi label $\mathscr{L}_{vor}$:

$$\mathscr{L}_{cluster/vor}(y,t) = -\frac{1}{|\Omega|}\sum_{i\in\Omega}\left[t_i \log y_i + (1-t_i)\log(1-y_i)\right], \tag{2}$$

where y is the probability map, t is the cluster label or Voronoi label, and Ω is the set consisting of non-ignored pixels. The final loss is $\mathscr{L}_{ce} = \mathscr{L}_{cluster} + \mathscr{L}_{vor}$.

2.3. Model refinement using dense CRF loss

The model trained using the two types of labels is able to predict the masks of individual nuclei with high accuracy. To further improve the performance, we refine the nuclear boundaries with the dense CRF loss. Previously post-processing such as region growing (Kumar et al., 2017), graph search (Yang et al., 2018) or dense CRF (Chen et al., 2015) is often utilized to refine the segmentation results. These algorithms introduce more computational complexity, making them unsuitable for the processing of large resolution Whole Slide Images. To solve this problem, similar to (Tang et al., 2018) we embed the dense CRF into the loss function to improve the accuracy. The loss function is not calculated during inference, and therefore will not introduce additional computational cost after training.

Let $\tilde{y} = (\tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_N)$ denote the predicted label (0 for background and 1 for nuclei) from probability maps y and t be the label. The dense CRF is to minimize the energy function:

$$E(\tilde{y},t) = \sum_i \phi(\tilde{y}_i, t_i) + \sum_{i,j} \psi(\tilde{y}_i, \tilde{y}_j), \tag{3}$$

where ϕ is the unary potential that measures how likely a pixel belongs to a certain class, and ψ is the pairwise potential that measures how different a pixel's label is from all other pixels' in the image. The unary term is replaced with the cross entropy loss $\mathscr{L}_{ce}$. The pairwise potential usually has the form:

$$\psi(\tilde{y}_i, \tilde{y}_j) = \mu(\tilde{y}_i, \tilde{y}_j) W_{ij} = \mu(\tilde{y}_i, \tilde{y}_j) \sum_{m=1}^{K} w_m k_m(\tilde{f}_i, \tilde{f}_j), \tag{4}$$

where μ is a label compatibility function, W_{ij} is the affinity between pixels i, j and is often calculated by the sum of Gaussian kernels k_m. In this work we choose μ as the Potts model, i.e., $\mu(\tilde{y}_i, \tilde{y}_j) = [\tilde{y}_i \neq \tilde{y}_j]$, and bilateral feature vector $\tilde{f}_i = \left(\frac{p_i}{\sigma_{pq}}, \frac{q_i}{\sigma_{pq}}, \frac{r_i}{\sigma_{rgb}}, \frac{g_i}{\sigma_{rgb}}, \frac{b_i}{\sigma_{rgb}}\right)$ that contains both location and color information. σ_{pq} and σ_{rgb} are Gaussian bandwidth.

To adapt the energy function to a loss function that is differentiable for training, we relax the pairwise potential as (Tang et al., 2018): $\psi(\tilde{y}_i, \tilde{y}_j) = \tilde{y}_i(1 - \tilde{y}_j) W_{ij}$. Therefore, the dense CRF loss can be expressed as:

$$\mathscr{L}_{crf}(y, t_{cluster}, t_{vor}) = \mathscr{L}_{ce}(y, t_{cluster}, t_{vor}) + \beta \mathscr{L}_{pair}(y), \tag{5}$$

where $\mathscr{L}_{pair}(y) = \sum_{i,j} y_i(1 - y_j) W_{ij}$ is the pairwise potential loss and β is the weighting factor. The CRF loss is used to fine-tune the trained model. Due to the large number of pixels in an image, the cost of directly computing the affinity matrix $W = [W_{ij}]$ is prohibitive. For instance, there are $N^2 = 1.6 \times 10^9$ elements in W for an image of size 200×200 that has $N = 40000$ pixels. We adopt fast mean-field inference based on high-dimensional filtering (Adams et al., 2010) to compute the pairwise potential part.

3. Experiments and Results

To validate our method, we apply it to two datasets of H&E stained histopathology images for nuclei segmentation and compare the results with fully supervised methods, including the same model trained with full masks, the CNN3 method proposed by Kumar et al. (Kumar et al., 2017) and the DIST method proposed by Naylor et al. (Naylor et al., 2018).

3.1. Datasets, evaluation and implementation details

Datasets The Lung Cancer dataset contains 40 images from 8 different lung cancer cases, and each case has 5 images of size about 900×900. These images are split into train, validation and test sets, consisting of 24, 8 and 8 images, respectively. Each set has at least one image of each case. Another dataset is publicly available, i.e., MultiOrgan dataset (Kumar et al., 2017). It consists of 30 image of size 1000×1000, which are taken from multiple hospitals and include a diversity of nuclear appearances from seven organs (Kumar et al., 2017). Both datasets have full mask annotation. We obtain the points annotation for the training sets by computing the central point of each nuclear mask.

Evaluation metrics Four metrics are used for evaluation, including pixel accuracy, pixel-level F1 score, object-level Dice coefficient (Sirinukunwattana et al., 2015) and the Aggregated Jaccard Index (AJI) (Kumar et al., 2017). The pixel-level F1 score is defined as $F1 = 2 \cdot TP/(2 \cdot TP + FP + FN)$, where TP, FP, FN are the numbers of true positive, false positive and false negative pixels,

Table 1: Results on Lung Cancer dataset using our methods in different settings.

Method	Pixel-level		Object-level	
	Acc	F1	Dice_{obj}	AJI
Full	*0.9615*	*0.8771*	*0.8521*	*0.6979*
Weak/Voronoi	0.9147	0.6596	0.6472	0.4791
Weak/Cluster	0.9188	0.7662	0.5936	0.2332
Weak w/o CRF	0.9413	0.8028	0.7885	0.6328
Weak w/ CRF	**0.9433**	**0.8120**	**0.8002**	**0.6503**

respectively. The object-level Dice coefficient is defined as

$$Dice_{obj}(\mathscr{G},\mathscr{S}) = \frac{1}{2}\left[\sum_{i=1}^{n_{\mathscr{G}}} \gamma_i Dice(G_i, S^*(G_i)) + \sum_{j=1}^{n_{\mathscr{S}}} \sigma_j Dice(G^*(S_j), S_j)\right] \tag{6}$$

where γ_i, σ_j are the weights related to object areas, $\mathscr{G}$, $\mathscr{S}$ are the set of ground-truth objects and segmented objects, $S^*(G_i)$, $G^*(S_i)$ are the segmented object that has maximum overlapping area with G_i and ground-truth object that has maximum overlapping area with S_i, respectively. The correspondence is built if the overlap area of two objects are more than 50%. This metric takes into account each object individually, and measures how well each segmented object overlaps with the ground truth objects, as well as how well each ground truth object overlaps the segmented objects (Sirinukunwattana et al., 2015). Another object-level metric AJI is proposed to evaluate the performance in nuclei segmentation and defined as

$$AJI = \frac{\sum_{i=1}^{n_{\mathscr{G}}} |G_i \cap S(G_i)|}{\sum_{i=1}^{n_{\mathscr{G}}} |G_i \cup S(G_i)| + \sum_{k \in K} |S_k|} \tag{7}$$

where $S(G_i)$ is the segmented object that has maximum overlap with G_i with regard to Jaccard index, K is the set containing segmentation objects that have not been assigned to any ground-truth object.

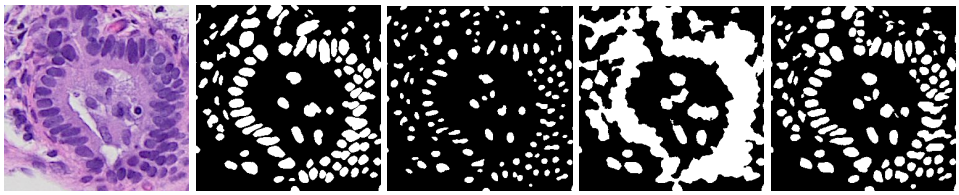
Implementation details Color normalization (Reinhard et al., 2001) is applied to all images to remove color variations caused by staining. Due to the small size of datasets, data augmentation such as random crop, scale, rotation, flipping, and affine transformation are adopted. The network is initialized with pretrained parameters and updated using the Adam optimizer. In weakly supervised settings, we train a model for 60 epochs with a learning rate of 1e-4, and fine-tune the model using dense CRF loss for 10 epochs with a learning rate of 1e-5. The parameters in CRF loss are $\sigma_{pq} = 10, \sigma_{rgb} = 10, \beta = 0.0005$. The validation set is not used because we have no access to ground-truth masks when training with points label. In fully supervised settings, we train 200 epochs using binary masks with a learning rate of 1e-4. The validation set is used to select the best model for test.

3.2. Results and comparison

The effects of two types of labels In order to show the importance of two types of generated labels, we report the results using either type of labels on the Lung Cancer datase in Table 1. Compared to the results using the cluster labels, those with Voronoi labels are better in the object-level

Table 2: Results on MultiOrgan dataset for CNN3 (Kumar et al., 2017), DIST (Naylor et al., 2017), fully supervised training and our methods with and without CRF loss.

Method	Pixel-level		Object-level	
	Acc	F1	$Dice_{obj}$	AJI
CNN3	-	-	-	0.5083
DIST	-	0.7623	-	**0.5598**
Full	*0.9194*	*0.8100*	*0.6763*	*0.3919*
Weak w/o CRF	0.9052	0.7745	0.7231	0.5045
Weak w/ CRF	**0.9071**	**0.7776**	**0.7270**	0.5097



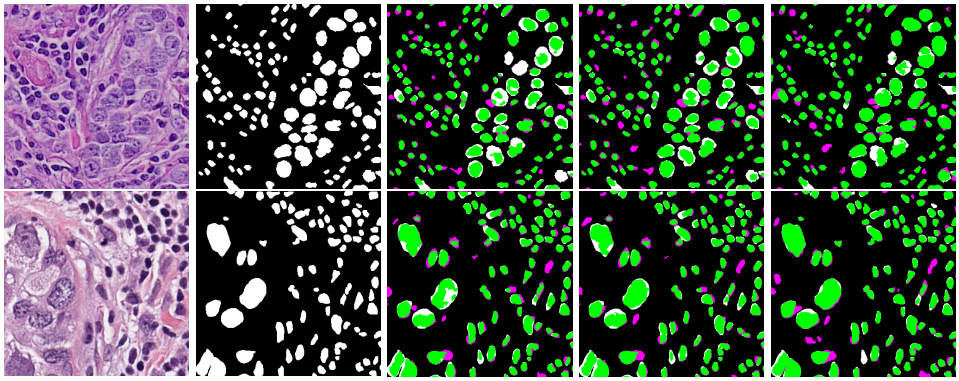
(a) image (b) gt mask (c) Voronoi labels (d) cluster labels (e) both labels

Figure 3: Results using different pixel-level labels: (a) image, (b) ground-truth mask, (c)-(e) are results using Voronoi labels, cluster labels and both labels, respectively.

metrics but worse in pixel-level metrics. This is because the model trained with Voronoi labels predicts the central parts of nuclei, resulting in small separated instances (Fig. 3(c)). While lacking the Voronoi edge information, the model using cluster labels is not able to separate close nuclei (Fig. 3(d)). In contrast, segmentation results using both labels are better than those with either label alone (Fig. 3(e)).

The effects of dense CRF loss From Table 1, it can be observed that the refinement with dense CRF loss improves the segmentation performance on the Lung Cancer dataset for all four metrics, but it is less effective on the MultiOrgan dataset. The reason is that in the MultiOrgan dataset there are many more crowded and touching nuclei that have no clear boundaries. CRF loss cannot handle these hard cases well.

Comparison to fully supervised methods The segmentation performance of our weakly supervised method is close to that of the fully supervised models with the same network structure. On the Lung Cancer dataset, the gaps for accuracy, F1 score, Dice and AJI are 1.9%, 7.4%, 6.1%, 6.8%, respectively. On the MultiOrgan dataset, the gaps for accuracy and F1 score are 1.3% and 4.0%. However, the fully supervised model has very low Dice and AJI, since for fair comparison we didn't perform post-processing to separate the touching nuclei for any of the methods. The weakly supervised model is able to separate most of them due to the Voronoi labels while the fully supervised model failed to achieve this. Compared to the CNN3 method in (Kumar et al., 2017), our method achieved the similar accuracy in terms of the AJI value. Compared to the state-of-the-art DIST method (Naylor et al., 2018), our approach has the higher pixel-level F1 score, but still has room



(a) images (b) gt masks (c) without CRF (d) with CRF (e) full annotation

Figure 4: Comparison of weakly and fully supervised training: (a) images, (b) ground-truth masks, (c)-(e) are results for weak labels without, with CRF loss and full labels, respectively, overlapped with ground-truth masks. Pixels in green, magenta, white are true positives, false positives and false negatives, respectively.

for improvement on the nuclear shapes, as indicated by the AJI values. Several image results are illustrated in Fig. 4.

Annotation time In order to show the time efficiency of points annotation, our pathologist annotated eight images (one per case) in the Lung Cancer dataset using points, bounding boxes and full masks, respectively. The average time spent on each image (about 600 nuclei in average) for full masks is 115 minutes while for bounding boxes, 67 minutes. However, it only takes about 14 minutes for points annotation.

4. Conclusion

In this paper we present a new weakly supervised nuclei segmentation method using only points annotation. We generate the Voronoi label and cluster label from the points label and take advantage of the dense CRF loss to refine our trained model. Our method is able to achieve comparable performance as fully supervised methods while requiring much less annotation effort which in turn allows us to analyze large amounts of data.

References

Andrew Adams, Jongmin Baek, and Myers Abraham Davis. Fast high-dimensional filtering using the permutohedral lattice. *Computer Graphics Forum*, 29(2):753–762, 2010.

Yousef Al-Kofahi, Wiem Lassoued, William Lee, and Badrinath Roysam. Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Transactions on Biomedical Engineering*, 57(4):841–852, 2010.

Amy Bearman, Olga Russakovsky, Vittorio Ferrari, and Li Fei-Fei. What's the point: Semantic segmentation with point supervision. In *European Conference on Computer Vision*, pages 549–565. Springer, 2016.

Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. Semantic image segmentation with deep convolutional nets and fully connected crfs. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

Jifeng Dai, Kaiming He, and Jian Sun. Boxsup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1635–1643, 2015.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

Andrew Janowczyk and Anant Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 7, 2016.

Hoel Kervadec, Jose Dolz, Meng Tang, Eric Granger, Yuri Boykov, and Ismail Ben Ayed. Constrained-cnn losses for weakly supervised segmentation. *Medical image analysis*, 54:88–99, 2019.

Anna Khoreva, Rodrigo Benenson, Jan Hendrik Hosang, Matthias Hein, and Bernt Schiele. Simple does it: Weakly supervised instance and semantic segmentation. In *CVPR*, page 3, 2017.

Henning Kost, André Homeyer, Jesper Molin, Claes Lundström, and Horst Karl Hahn. Training nuclei detection algorithms with simple annotations. *Journal of Pathology Informatics*, 8, 2017.

Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging*, 36(7):1550–1560, 2017.

Di Lin, Jifeng Dai, Jiaya Jia, Kaiming He, and Jian Sun. Scribblesup: Scribble-supervised convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3159–3167, 2016.

Faisal Mahmood, Daniel Borders, Richard Chen, Gregory N McKay, Kevan J Salimian, Alexander Baras, and Nicholas J Durr. Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *arXiv preprint arXiv:1810.00236*, 2018.

Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Nuclei segmentation in histopathology images using deep neural networks. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pages 933–936. IEEE, 2017.

Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Transactions on Medical Imaging*, 2018.

George Papandreou, Liang-Chieh Chen, Kevin P Murphy, and Alan L Yuille. Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation. In *Proceedings of the IEEE international conference on computer vision*, pages 1742–1750, 2015.

Deepak Pathak, Philipp Krahenbuhl, and Trevor Darrell. Constrained convolutional neural networks for weakly supervised segmentation. In *Proceedings of the IEEE international conference on computer vision*, pages 1796–1804, 2015.

Hui Qu, Gregory Riedlinger, Pengxiang Wu, Qiaoying Huang, Jingru Yi, Subhajyoti De, and Dimitris Metaxas. Joint segmentation and fine-grained classification of nuclei in histopathology images. In *International Symposium on Biomedical Imaging*, pages 900–904. IEEE, 2019.

Martin Rajchl, Matthew CH Lee, Ozan Oktay, Konstantinos Kamnitsas, Jonathan Passerat-Palmbach, Wenjia Bai, Mellisa Damodaram, Mary A Rutherford, Joseph V Hajnal, Bernhard Kainz, et al. Deepcut: Object segmentation from bounding box annotations using convolutional neural networks. *IEEE transactions on medical imaging*, 36(2):674–683, 2017.

Erik Reinhard, Michael Adhikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Computer Graphics and Applications*, 21(5):34–41, 2001.

Olaf Ronneberger, Philipp Fischer, and Thomas 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.

Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.

Korsuk Sirinukunwattana, David RJ Snead, and Nasir M Rajpoot. A stochastic polygons model for glandular structures in colon histology images. *IEEE transactions on Medical Imaging*, 34(11): 2366–2378, 2015.

Meng Tang, Federico Perazzi, Abdelaziz Djelouah, Ismail Ben Ayed, Christopher Schroers, and Yuri Boykov. On regularized losses for weakly-supervised cnn segmentation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 507–522, 2018.

Mitko Veta, Paul J Van Diest, Robert Kornegoor, André Huisman, Max A Viergever, and Josien PW Pluim. Automatic nuclei segmentation in h&e stained breast cancer histopathology images. *PloS one*, 8(7):e70221, 2013.

Fuyong Xing, Yuanpu Xie, and Lin Yang. An automatic learning-based framework for robust nucleus segmentation. *IEEE transactions on Medical Imaging*, 35(2):550–566, 2016.

Lin Yang, Yizhe Zhang, Zhuo Zhao, Hao Zheng, Peixian Liang, Michael TC Ying, Anil T Ahuja, and Danny Z Chen. Boxnet: Deep learning based biomedical image segmentation using boxes only annotation. *arXiv preprint arXiv:1806.00593*, 2018.

Zhuo Zhao, Lin Yang, Hao Zheng, Ian H Guldner, Siyuan Zhang, and Danny Z Chen. Deep learning based instance segmentation in 3d biomedical images using weak annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 352–360. Springer, 2018.