Attentional Dilated Convolution Neural Network for Nuclei Segmentation in Histopathology Images

Chao Yi
Dept. of Instrument Science and Engineering
Shanghai Jiao Tong University
Shanghai , China
ucnpliterme@sjtu.edu.cn

Xuelei Chen
Dept. of Instrument Science and Engineering
Shanghai Jiao Tong University
Shanghai , China
chenxuelei@sjtu.edu.cn

Lingwei Quan
Dept. of Instrument Science and Engineering
Shanghai Jiao Tong University
Shanghai , China
caleenkwon@sjtu.edu.cn

Cunyue Lu*
Dept. of Instrument Science and Engineering
Shanghai Jiao Tong University
Shanghai , China
lucunyue@sjtu.edu.cn

***Abstract*—The accurate nuclei segmentation in histopathology images is the foundational task of pathological diagnosis. Deep learning has emerged as an effective method of nuclei segmentation, but the touching or overlapping nuclei are difficult to segment. In this paper, we propose a novel method for nuclei segmentation. The proposed method uses channel attention mechanism to reweight channel-wise features and further improve segmentation performance. Moreover, dilated convolution is used to aggregate multi-scale contextual information and preserve the original resolution. We name our method as Attentional Dilated Convolution Neural Network (ACDNN). Experiments on a large tissue image dataset show the advantage of our method over other state-of-the-art methods.**

Keywords—nuclei segmentation, channel attention, dilated convolution, histopathology images

I. INTRODUCTION

Histopathology image analysis plays a critical role in cancer detection, diagnosis and other clinical analyses [1][5]. Nuclei segmentation is the fundamental task of pathological image analysis because the morphology and distribution of the nuclei are closely related to the presence and severity of diseases [1]. Experienced pathologists usually segment histopathology images manually, which is tedious and subjective. Therefore, segmentation methods for efficient and objective analysis of histopathological images have been studied [2]-[4].

Traditional nuclei segmentation methods include thresholding, morphological operations, watershed and Level Sets [5]. Thresholding and morphological operations are long-established basic image processing methods, and the principle and implementation are relatively simple. However, in the face of complex image background, satisfactory segmentation results are often difficult to obtain [6][7]. Watershed method works well when the position of the front object and the background can be marked [8]. Level Sets are specifically designed to handle topological changes, but they are not robust to boundary events [9]. Although these studies have conducted profound discussions on nuclei segmentation, there are still problems such as low accuracy or poor precision. Therefore, relevant studies have begun to turn their attention to deep learning algorithms.

Convolutional Neural Network (CNN) is a type of feedforward neural network that includes convolution calculations and has a deep structure. It is one of the representative algorithms of deep learning. In recent years, the convolutional neural network model has reached or even surpassed the human level in a large number of applications in the field of computer vision, so the CNN model has gradually become the first choice for numerous image applications. Current advanced nuclei segmentation methods apply full convolutional network (FCN) and their derived version [10][11], which use encoders to extract features and restore high-resolution predictions from low-resolution feature maps. Compared with traditional methods, they have achieved good segmentation results. But the down-sampling in these methods may reduce the positioning accuracy of the boundary in nuclei segmentation. Qu et al. [12] proposed a full-resolution neural network (FullNet) to improve positioning accuracy and achieved better segmentation results. However, the network treats all channels of feature maps equally, which limits the network's discrimination learning ability to a certain extent.

Therefore, we propose an attentional dilated convolution neural network using channel attention mechanism to further improve the segmentation performance [13]. The essence of the channel attention mechanism is to model the importance among various features and to suppress useless information, which is simple and effective. Moreover, we adopt dilated convolution to preserve the original resolution by aggregating multi-scale contextual information [14].

Overall, the contribution of this paper is to propose a novel network architecture ADCNN, and sufficient experiments have proved the effectiveness of this method.

II. METHODS

A. ADCNN Architecture

The architecture of our ADCNN is shown in Fig.1. It consists of a 3×3 convolutional layer, seven attentional dilated dense blocks with different dilation factors, and a final 3×3 convolutional layer. The structure of the attentional dilated dense block is shown in Fig.2. It consists of a dense block, a 1×1 convolutional layer and a channel attention block.

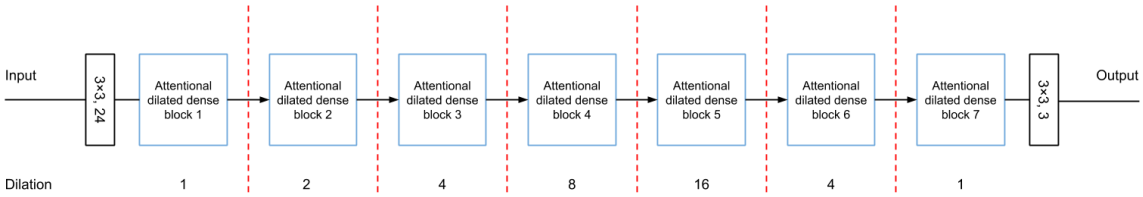

Fig. 1. The architecture of ADCNN. Each red dash line represents a change of dilation factor in ADCNN

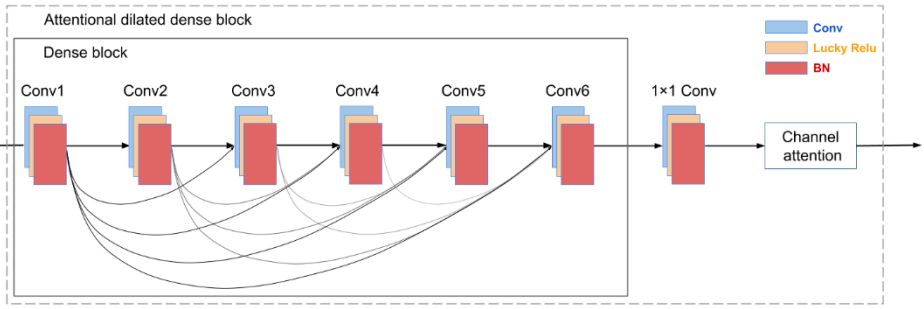

Fig. 2. The structure of the attentional dilated dense block.

We adopt the dense block structure to reduce the number of parameters, which effectively avoid over-fitting, and require less computation. In one dense block, each layer uses all preceding feature maps as input, and its feature-maps are used as inputs into all subsequent layers. This structure can strengthen feature propagation and encourage feature reuse, thereby achieving good performance with fewer parameters [15]. We adopt Conv-LeakyReLU-BN structure in dense blocks to reduce the memory requirements of intermediate feature maps [12]. The 1×1 convolutional layer after each dense block can further compress the number of output channels by half, reducing the amount of calculation.

In the proposed ADCNN, we use dilated convolution in dense blocks to obtain the large receptive field by aggregating multi-scale contextual information. We set the dilation factors as (1,2,4,8,16,4,1) by referring to the FullNet [12]. Dilation 4 and 1 are added after block 5 to improve the gridding artifacts resulted from large dilation factors. The hybrid dilations in each dense block are presented in Table I, corresponding to six different convolutional layers.

The channel attention mechanism in Fig.3 is adopted in our ADCNN, and its main function is to re-weight channel-wise features. It consists of three parts: squeeze, excitation, and attention. First, global average pooling is used to squeeze the feature map in the channel attention branch. Then two convolutional layers are used to calculate weights for channels. The output of the channel attention branch and the original channels are multiplied to generate a re-weighted feature map. This mechanism improves the representation ability of the feature map, thus improving image segmentation performance.

TABLE I. DILATION RATES IN DIFFERENT CONVOLUTION LAYERS

Factor	Conv1	Conv2	Conv3	Conv4	Conv5	Conv6
1	1	1	1	1	1	1
2	1	2	3	1	2	3
4	1	2	3	5	6	7
8	2	5	7	9	11	14
16	10	13	16	17	19	21
4	1	2	3	5	6	7
1	1	1	1	1	1	1

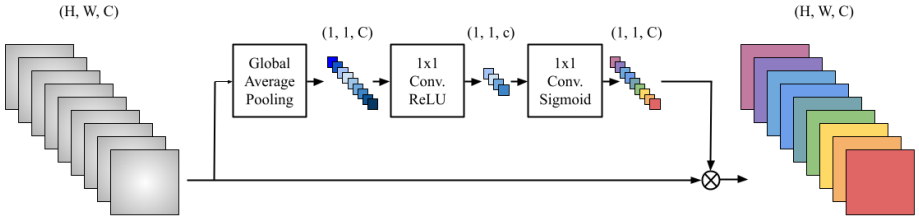

Fig. 3. Illustration of channel attention mechanism.

B. Loss Function

To demonstrate the effectiveness of the proposed attention dilated learning method, we used the same loss function as that used in FullNet [12]. Cross entropy loss realizes the correct classification of each pixel, but lacks the spatial relationship between pixels, which limits the nuclei segmentation performance. The variance constrained cross entropy (varCE) overcomes this limitation by learning the spatial relationship between pixels in the same instance to improve localization accuracy.

The cross entropy loss of class M is formulated as

$$\mathcal{L}_{CE}(y,t,w) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{M} w_i t_i^{(m)} log y_i^{(m)} \quad (1)$$

where $y_i^{(m)}$ denotes the probability of pixel i belonging to class M, N is the number of all pixels, $t_i^{(m)}$ denotes the groundtruth label for class M, w_i is the weight of pixel i.

The variance term loss is formulated as

$$\mathcal{L}_{var}(y,t) = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{|S_c|}\sum_{i=1}^{|S_c|}(\mu_c - \hat{y}_i)^2 \quad (2)$$

where C is the number of instances, S_c is the pixels set belonging to instance c, $\hat{y}_i$ denotes the probability of the correct class for pixel i, μ_c is the mean value of $\hat{y}_i$ in set S_c.

Combining two loss functions mentioned above, the final varCE is formulated as

$$\mathcal{L}_{varCE} = \mathcal{L}_{CE} + \alpha\mathcal{L}_{var} \quad (3)$$

where α means the the weight of the variance term in the loss function.

III. Experiments

A. Implementation Details

The experiments are conducted on a large publicly accessible dataset of H&E stained tissue images with more than 21,000 painstakingly annotated nuclear boundaries, whose quality was validated by pathologist doctor [16]. To show the effectiveness of the proposed ADCNN, we use the same evaluation metrics adopted in the dataset paper [16], including F1-score, average Dice coefficient, average Hausdorff distance, and the Aggregated Jaccard Index (AJI). Our codes are trained and tested on Google Colab using Tesla P100 GPU 12GB. Taking into account the limitation of GPU memory, the batch size is set to 6. The weight α in the loss function is set to 1, the learning rate and training epoch are set to 0.001 and 300. Data augmentations like random scale, flip, rotation, affine and elastic transformation are adopted in the proposed ADCNN.

B. Data Preparation

Before the training process, training images and labels are supposed to be prepared. Ternary labels and weight maps are used to help separate close instances. Fig.4 shows two examples of (image, ternary label, weight map) pairs in the H&E stained tissue image dataset. Ternary label consists of three parts: inside areas, edges and background areas. The weight map is to encourage the network to learn the background between the touching instances. The background between the touching instances has a high weight in the loss function. The data preparation step is necessary because the morphology and distribution of the histopathology images are various and complicated.

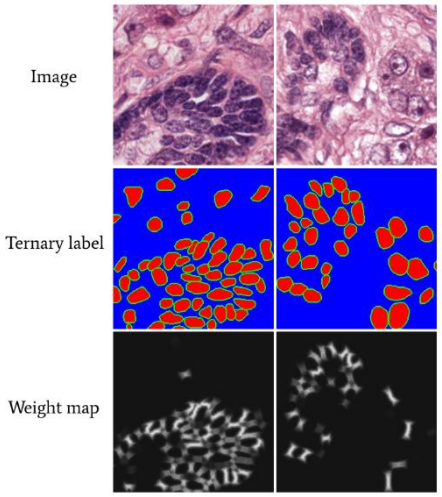

Fig. 4. Two examples of (image, ternary label, weight map) pairs

C. Results and Comparison

We compare the ADCNN with the state-of-the-art methods including FCN-Pooling [12], FullNet [12] and CNN3 [16]. Quantitative and qualitative results are presented in Table II and Fig.5 respectively. Bold numbers in Table II indicate that the corresponding method has achieved the best performance on the corresponding evaluation metrics. The quantitative results in Table II show the advantage of our method over other methods in all evaluation metrics.

Different colors in the image represent different instances. We use rectangles to mark the contrasting parts where our method is better than other methods in nuclei segmentation performance. Referring to the ground-truth, the proposed method successfully segmented two close instances while FCN-Pooling and FullNet did not in the rectangles.

TABLE II. Performance comparison with different methods

Method	F1	Dice	H	AJI
CNN3 [16]	0.8267	0.7623	7.6615	0.5083
FCN-Pooling[12]	0.8300	0.7595	10.1168	0.4783
FullNet[12]	0.8514	0.7875	7.9232	0.5467
ADCNN	**0.8608**	**0.7941**	**7.3389**	**0.5694**

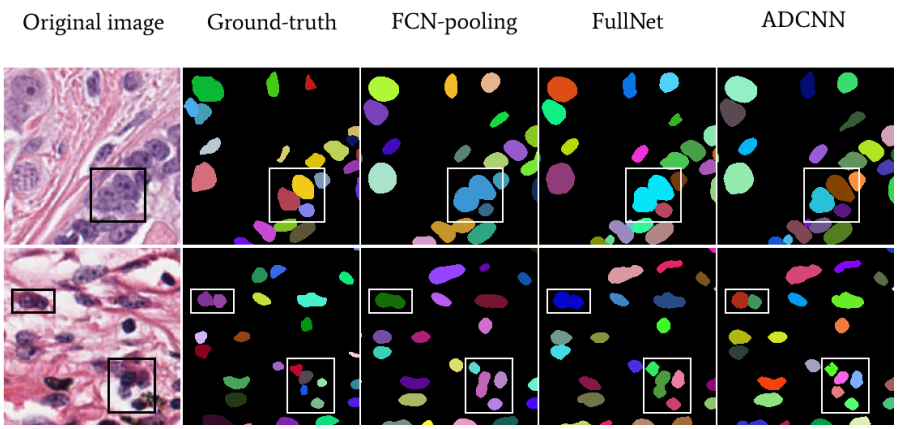

Fig. 5. Qualitative results comparison of nuclei segmentation

IV. Conclusion and Future Work

In this paper, we present an attentional dilated convolution neural network for nuclei segmentation in histopathology images. We introduce channel attention mechanism to this work and propose an attentional dilated dense block to build the network. Experiments on a large tissue image dataset show the effectiveness of the proposed method. The method in this paper achieves great improvement over other state-of-the-art methods. Future work can be focused on researching algorithms that can be applied to more complex cell environments. Another research direction is to reduce the computational cost while maintaining nuclei segmentation performance.

References

[1] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot and B. Yener, "Histopathological Image Analysis: A Review," *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 147-171, Oct 2009

[2] D. N. Louis, M. Feldman, A. B. Carter, S. Dighe, J. D. Pfeifer and B. Lynn et al., "Computational Pathology: A Path Ahead," Archives of Pathology & Laboratory Medicine, vol. 140, no. 1, pp. 41-50, Jan 2016

[3] H. Qu, J. Yi, Q. Huang, P. Wu and D. Metaxas, "Nuclei Segmentation Using Mixed Points and Masks Selected From Uncertainty," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, May 2020, pp. 973-976

[4] P. Naylor, M. Lae, F. Reyal and T, walter, "Segmentation of Nuclei in Histopathology Images by Deep Regression of the Distance Map." IEEE Transactions on Medical Imaging, vol. 38, no. 2, pp. 1-1, Aug 2018.

[5] T. Hayakawa, V. B. S. Prasath, H. Kawanaka, B. J. Aronow and S. Tsuruoka,"Computational Nuclei Segmentation Methods in Digital Pathology: A Survey," Archives of Computational Methods in Engineering, pp. 1-13, Sept 2019

[6] N. J. Gadgil, P. Salama, K. W. Dunn and E. J. Delp, "Nuclei segmentation of fluorescence microscopy images based on midpoint analysis and marked point process," 2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), Santa Fe, NM, Mar 2016, pp. 37-40

[7] K. Y. Win and S. Choomchuay, "Automated segmentation of cell nuclei in cytology pleural fluid images using OTSU thresholding," 2017 International Conference on Digital Arts, Media and Technology (ICDAMT), Chiang Mai, Mar 2017, pp. 14-18

[8] M. Veta, P. J. Diest and R. Kornegoor, A. Huisman, M. A. Viergever, J. P. W. Pluim, "Automatic nuclei segmentation in H&E stained breast cancer histopathology images." PloS one, vol. 8, no. 7, pp. e70221, Jul 2013

[9] X. Qi, F. Xing, D. J. Foran and L. Yang, "Robust Segmentation of Overlapping Cells in Histopathology Specimens Using Parallel Seed Detection and Repulsive Level Set," IEEE Transactions on Biomedical Engineering, vol. 59, no. 3, pp. 754-765, Mar 2012

[10] H. Chen, X. Qi, L. Yu and P. Heng, "DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, Dec 2016, pp. 2487-2496

[11] X. Xu, Q. Lu, L. Yang, X. Hu, D. Chen and Y. Hu et al., "Quantization of Fully Convolutional Networks for Accurate Biomedical Image Segmentation," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, Dec 2018, pp. 8300-8308

[12] H. Qu, Z. Yan, G. M. Riedlinger, S. De, and D. N. Metaxas, "Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss," International Conference on Medical Image Computing and Computer-Assisted Intervention, Oct 2019, pp. 378-386

[13] J. Hu, L. Shen, S. Albanie, G. Sun and E. Wu, "Squeeze-and-Excitation Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 8, pp. 2011-2023, Aug. 2020

[14] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," arXiv preprint arXiv:1511.07122, Nov 2015

[15] G. Huang, Z. Liu, L. V. D. Maaten and K. Q. Weinberger, "Densely connected convolutional networks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Aug 2017, pp. 4700-4708

[16] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane and A. Sethi, "A Dataset and a Technique for Generalized Nuclear Segmentation for Computational Pathology," IEEE Transactions on Medical Imaging, vol. 36, no. 7, pp. 1550-1560, July 2017