

# Integrating deep convolutional neural networks with marker-controlled watershed for overlapping nuclei segmentation in histopathology images



Lipeng Xie, Jin Qi\*, Lili Pan, Samad Wali

School of Information and Communication Engineering, University of Electronic Science and Technology of China, China

## ARTICLE INFO

### Article history:

Received 10 March 2019  
Revised 24 September 2019  
Accepted 26 September 2019  
Available online 10 October 2019

Communicated by Pingkun Yan

### Keywords:

Histopathology images  
Overlapping nuclei segmentation  
Convolutional neural networks  
Marker extraction  
Watershed transformation

## ABSTRACT

Nuclei segmentation in histopathology images plays a crucial role in the morphological quantitative analysis of tissue structure and has become a hot research topic. Though numerous efforts have been tried in this research area, the overlapping and touching nuclei segmentation remains a challenging problem. In this paper, we present a novel and effective instance segmentation method for tackling this challenge by integrating Deep Convolutional Neural Networks with Marker-controlled Watershed. Firstly, we design a novel network architecture with multiple segmentation tasks, called Deep Interval-Marker-Aware Network, for learning the foreground, marker, and interval of nuclei, simultaneously. Then the learned interval between overlapping nuclei is used to refine the foreground result of nuclei by using the logical operators. Finally, the learned marker result and the nuclei segmentation result refined by interval are transmitted into the Marker-controlled Watershed for splitting the touching nuclei. The experiments on the standard public datasets demonstrate that our method achieves a substantial improvement compared with state-of-the-art methods. Source codes are available at [https://github.com/appiek/Nuclei\\_Segmentation\\_Experiments\\_Demo](https://github.com/appiek/Nuclei_Segmentation_Experiments_Demo).

© 2019 Elsevier B.V. All rights reserved.

## 1. Introduction

Histopathology is about the investigation of the manifestations of disease based on microscopic examination of tissue structure, which has been viewed as a gold standard for the diagnosis of most cancer diseases, such as the breast cancer [1], lung cancer [2], and prostate cancer [3]. In general, the tissue specimens of patients should be stained with special colors for highlighting the interesting components of tissue such as the lymphocytes, cancer nuclei, and glands [4]. For example, hematoxylin & eosin (H&E) is one of the most widely used stain materials, which imparts the blue-purple and pink color to the nuclei and cytoplasm respectively (see Fig. 1). Then the pathologists would observe the characteristics of histological structure under a microscope, which is an important step for grading and diagnosis of disease [5,6]. Nevertheless, this work is very laborious and time-consuming for pathologists. In addition, the diagnosis result is likely to be influenced by the cognitive diversity due to the lack of quantitative analysis. With the development of digital pathology technology, the image information of tissue can be captured by the whole slide imaging (WSI)

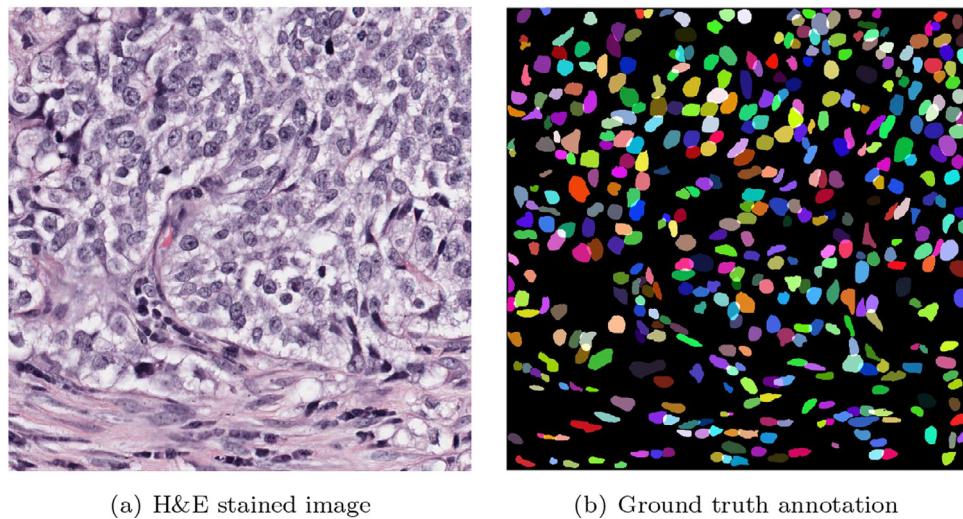
scanners, stored as digital image by compression algorithm and displayed in the computer screen [7]. Furthermore, some researchers have employed image processing and artificial intelligence technology to establish computer-aided diagnosis tools [8–10] for improving the work efficiency and reducing error rates for pathologists.

Segmentation of cell nuclei is an indispensable step for automatic digitized histopathology imagery analysis system [11,12]. Unfortunately, as stated in literatures [11,13–16], this task is difficult due to (1) the dense overlapping between nuclei, (2) the complex variability in size, shape, appearance, and texture of the individual nuclei, and (3) the non-homogenous background, as shown in Fig. 1. In addition, the difference in tissue type, cell type and staining material of histopathology images is also a great challenge to the robustness of automatic segmentation algorithms [12,17]. In this paper, we mainly focus on the overlapping nuclei segmentation task, which is to divide the touching nuclei into several individual nucleus instances.

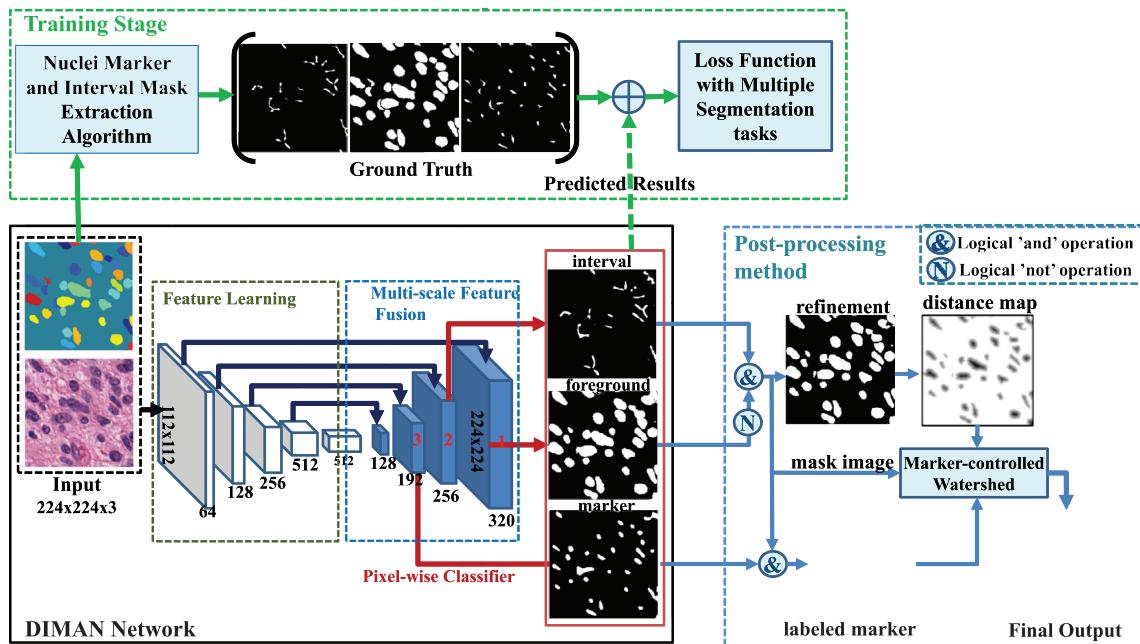
To solve this problem, we propose an efficient computing framework by combining marker-controlled watershed (MCW) [18] with convolutional neural networks (CNN) [19], as illustrated in Fig. 2. The main novelty and contribution of our work can be summarized as below: (1) presenting a multi-task network architecture to simultaneously learn the foreground, interval,

\* Corresponding author.

E-mail address: [jqi@uestc.edu.cn](mailto:jqi@uestc.edu.cn) (J. Qi).



**Fig. 1.** The example of overlapping nuclei segmentation. The original image (left side) and the associated segmentation ground truth (right side).



**Fig. 2.** The framework of our proposed method.

and marker information of nuclei from the original images, (2) integrating the marker and interval results with the MCW method as a post-processing method for separating the overlapping nuclei in the foreground, and (3) proposing an efficient method to extract the mask of nuclei interval and marker from the annotation of nuclei foreground for training our network. Note that the interval of nuclei indicates the gap between adjacent nuclei, and the marker of nuclei represents the central location of each nucleus. These important prior information can greatly improve the robustness and accuracy of MCW method in touching nuclei segmentation. By contrast, Naylor et al. [17,20] simply extracted the mark of nuclei from the output of CNN with ignoring the importance of interval information. Moreover, the errors in segmentation result would impair the accuracy of mark extraction method, and further lead to the over-segmentation and under-segmentation problems for MCW method. The comparison with state-of-the-art methods can fully prove the superiority and effectiveness of our method.

The remainder of our paper is organized as follows. In [Section 2](#), we review some related work with overlapping nuclei segmentation. [Section 3](#) introduces the framework of our method. The implementation details of our method are presented in [Section 4](#). Then the experimental results are shown in [Section 5](#). Finally, the conclusion and perspectives are drawn in [Section 6](#).

## 2. Related work

Thresholding and morphological operation were two widely used nuclei segmentation methods in the early stage of this research area [21]. For examples, Gurcan et al. [22] established a nuclei segmentation framework for neuroblastoma cancer analysis by combining morphological reconstruction with hysteresis thresholding [23]. Nawandhar et al. [24] adopted the Otsu method [25] and multiple morphological operations to segment the cell nuclei in colon tissue. In addition, Huang et al. [26] proposed a multilevel Otsu thresholding method for leukocyte nucleus

segmentation. However, due to the significant variations of pixel intensity, the contrast between the foreground and background in histopathology images is not distinct. This problem has great negative effect on the nuclei segmentation performance for the thresholding methods.

Watershed transform is a common solution for overlapping objects segmentation. In [27], the Otsu algorithm [25] was used to convert the nuclei image to binary segmentation result. Then the classical watershed transform was employed to isolate cluster nuclei from each other by modeling the input grayscale image as a topographic surface and flooding from the regional minima to establish the watershed lines. Nevertheless, one nucleus instance often contains multiple regional minima due to the inhomogeneous distribution of pixel intensity, which leads to the over-segmentation. For solving this issue, Yang et al. [28] proposed an efficient marker extraction method based on the conditional erosion. The nuclei foreground result would be processed by the conditional erosion twice with coarse and fine erosion structure successively, then the reserved foreground would be utilized as the markers of nuclei. Naylor et al. [17] detected the local maxima of the posterior probability of nuclei foreground as the seed points of nuclei, and presented a criterion to decide whether or not to split the connected region based on the morphological dynamics.

Moreover, the advanced image segmentation methods such as the deformable models and graph-cut method were also employed to segment the overlapping nuclei with obtaining good performance. For instance, Fatakdawala et al. [29] presented a lymphocyte segmentation framework by combining the geodesic active contour model with expectation-maximization (EM) algorithm. Yan et al. [30] proposed a multiphase level set approach for delineating the boundaries of cell. In literature [31], the local characteristics and prior knowledge of the nuclei boundary were incorporated into the active contour model. Similarly, Qi et al. [12] utilized the mean-shift clustering method to detect the center location of nuclei as the initialization of level set approach. In literature [13], an improved laplacian-of-gaussian filtering algorithm was presented to detect the seed points of nuclei, and then the graph-cut method was used to segment the cell nuclei images. However, the performance of above nuclei image segmentation techniques are sensitive to the initialization results and the setting of parameters.

Recently, numerous CNN-based segmentation methods have been proposed for nuclei images with achieving outstanding performance [11,17,32,33], in which all pixels are assigned a specific label (e.g., the value of 0 and 1 denote the foreground and background respectively). For example, Xing et al. [11] applied the deep CNN to extract the foreground result of nuclei, and presented a shape deformation method to separate the touching nuclei. Kumar et al. [32] presented a generalized nuclei segmentation technique based on CNN and region growing. In their approach, the CNN was applied to identify the foreground and boundary of nuclei, and then the seeds of nuclei would grow toward the boundary under the predefined criterion for separating the overlapping nuclei. Though this approach obtained good results, it is time-consuming with around 30 s for each  $1000 \times 1000$  image. In [33], Chen et al. proposed a novel CNN with multi-task learning and pixels-to-pixels characteristics for extracting the foreground and boundary of nuclei at the same time. Then the touching nuclei were divided into individual ones by removing the boundary from the foreground. Nevertheless, the de-boundary operation would destroy the shape information of nuclei. The combination of CNN and watershed method was widely applied in overlapping nuclei segmentation. Naylor et al. [17] employed the CNN to segment the foreground of nuclei, and presented a watershed transformation based post-processing method to segment the nuclei in the foreground. In their further work [20], they transformed the nuclei segmentation problem into a regression task by using CNN

to predict the distance map of nuclei and utilize the same post-processing method to refine the result.

### 3. Our proposed method

In this paper, we design a novel deep interval-marker-aware network (DIMAN) with end-to-end training and pixel-wise annotation properties. In our network, the multi-scale feature maps are stacked as a feature pyramid by skip connection. Moreover, three pixel-wise classifiers are connected with the feature pyramid in parallel for learning multiple classes of objects from the original images. Then we systematically combine these learned results into MCW method to obtain the final segmented and separated nuclei.

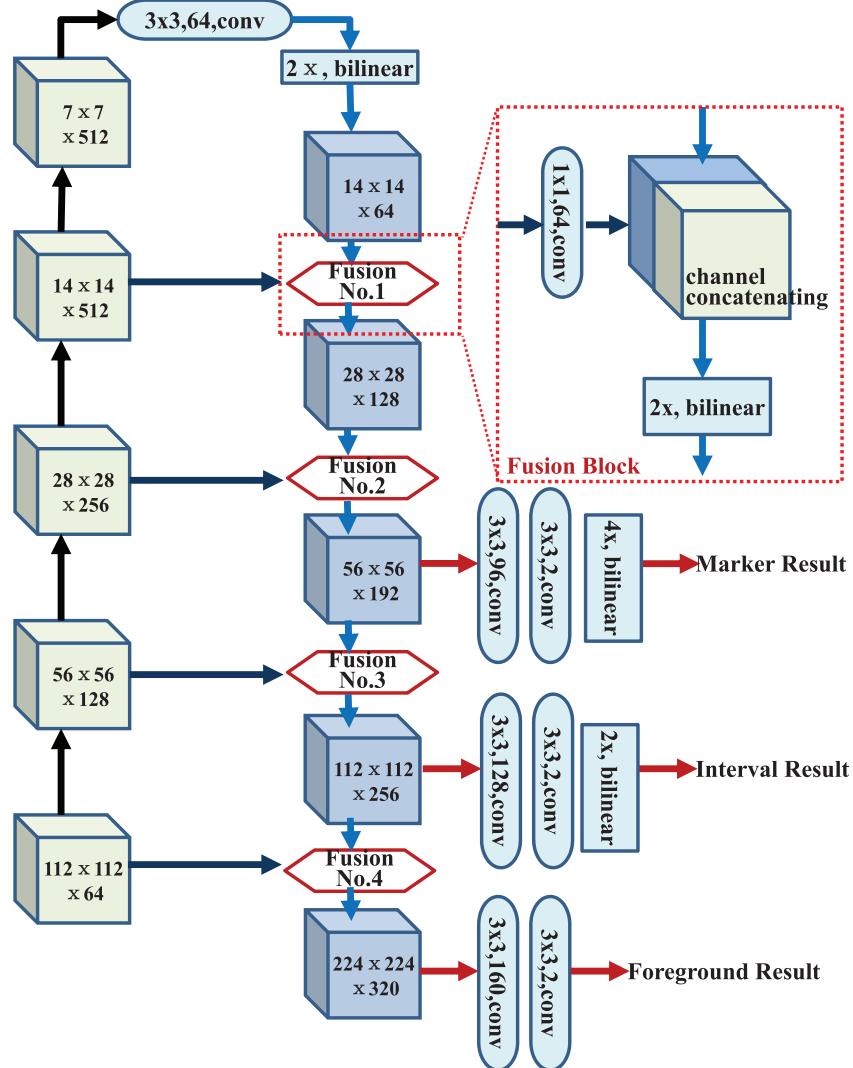
#### 3.1. Network architecture

As illustrated in Fig. 2, our network is mainly composed of five parts: a feature learning network for extracting the hierarchical features from input images, a multi-scale feature fusion network for upsampling and merging the hierarchical features as a feature pyramid, and three pixel-wise classifiers (denoted by thick red arrow) for segmenting the foreground, interval, and marker of nuclei synchronously. To reduce the mutual interference among the multiple tasks, we employ different feature maps to predict different classes of objects in the original input images. For segmenting the marker of nuclei, we utilize the feature map in third-layer (high-level), for interval we use the feature map in second-layer (middle-level), and for the foreground we take advantage of the feature map in first-layer (low-level). The main reason for designing this structure is that the segmentation of foreground is heavily depend on the edge information of image, which mainly exists in the low-level feature map. Nevertheless, the object-level information existing in the high-level feature map is advantageous to the prediction of nuclei marker.

In Fig. 3, we show the detailed structure information of our network. The architecture of feature learning network is same with the first thirteen convolutional layers of VGG16 network [34]. This architecture can transform the original input image into five different feature maps including high-level and low-level contextual information by forward propagation. However, the resolution of feature maps would be reduced 32 times than the original image. In order to restore the spatial information of feature map and merge the hierarchical features as a feature pyramid, we design a multi-scale feature fusion network, which consists of four fusion blocks. As illustrated in the red dotted rectangle of Fig. 3, each fusion block is made up of a  $3 \times 3$  convolutional operation with filter size 1 and filter number 64 for reducing the channel number of feature maps from the VGG16 network, a concatenating operation for merging the feature maps along the channel axis, and a bilinear interpolation operation with factor 2 for upsampling the merged feature map. The four fusion blocks can stack the multi-scale feature maps as a feature pyramid, which contains both the pixel-level and object-level information. Since the bilinear interpolation operation has no parameters needed to be optimized, it can reduce the model complexity degree and the training time of our network. Each pixel-wise classifiers consists of two duplicate  $3 \times 3$  convolutional operations with stride 1. Especially, the pixel-wise classifiers of interval and marker are connected with a bilinear interpolation operation since the resolution of feature maps is lower than the original image. All the classifiers would make contribution to the loss function for training the proposed network.

#### 3.2. Loss function with multiple segmentation tasks

The common way of training CNN model is to minimize the loss function of CNN by back-propagating algorithm. Given a train-



**Fig. 3.** The structure of our network. The white cube represents the feature map learned by VGG16, and the blue cube indicates the merged feature map by fusion block.

ing dataset with  $N$  image patches denoted by  $\{I^i(x) \in \mathbb{R}, G^i(x) \in \{0, 1\} | i = 1 \dots N, x \in \Omega\}$ , where  $I$ ,  $G$ , and  $\Omega$  represent the input raw images with  $T$  pixels, the ground truth of image segmentation, and the 2-dimensional spatial domain of image respectively. We also utilize  $O^i(x) \in \mathbb{R}^2$  to denote the output of pixel-wise classifier. Then the output would be fed into the softmax function for predicting the label of pixel. For example, the probability of pixel  $I^i(x)$  attribute to label  $l$  is given by:

$$P(y = l | I^i(x)) = \frac{\exp(O^{ij}(x, l))}{\sum_{l'=0}^1 \exp(O^i(x, l'))} \quad (1)$$

Based on the probability map, the cross-entropy loss function of pixel-wise classification can be defined as below:

$$L(I, G; W) = -\frac{1}{N \times T} \sum_{i=1}^N \sum_{x \in \Omega} \sum_{l=0}^1 1\{G^i(x) = l\} \log(P(y = l | I^i(x))) \quad (2)$$

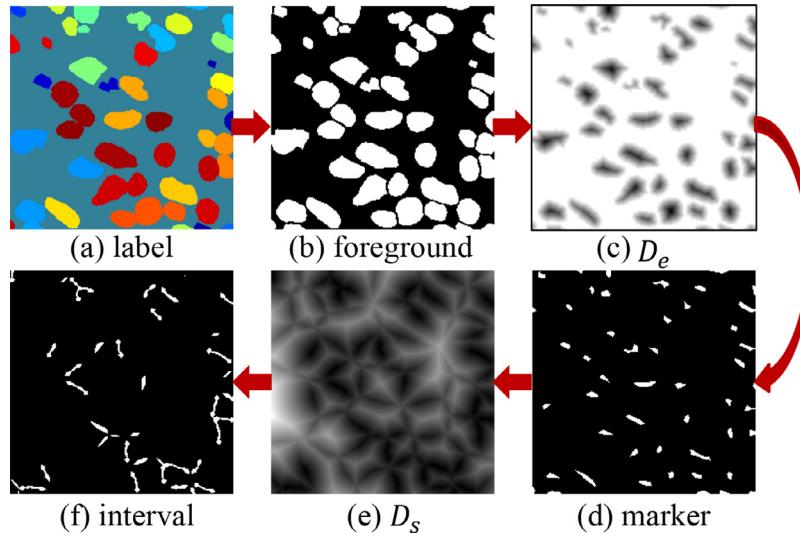
where  $W$  indicates the weight values associated with the connection between units in neural network, and  $1\{\cdot\}$  denotes the indicator function.

Actually, the ground truth data  $G^i(x)$  in our method contains three kinds of pixel-wise label: the foreground label  $G_F^i(x) \in \{0, 1\}$ ,

the interval label  $G_I^i(x) \in \{0, 1\}$ , and the marker label  $G_M^i(x) \in \{0, 1\}$ . Thus the loss function of our network should be composed of three parts:

$$\begin{aligned} L(I, G; W) = & -\frac{1}{N \times T} \sum_{i=1}^N \sum_{x \in \Omega} \\ & \times \left[ c_F \sum_{l_F=0}^1 1\{G_F^i(x) = l_F\} \log(P_F(y = l_F | I^i(x))) w_F(x) \right. \\ & + c_M \sum_{l_M=0}^1 1\{G_M^i(x) = l_M\} \log(P_M(y = l_M | I^i(x))) w_M(x) \\ & \left. + c_I \sum_{l_I=0}^1 1\{G_I^i(x) = l_I\} \log(P_I(y = l_I | I^i(x))) w_I(x) \right] \quad (3) \end{aligned}$$

where  $c_F$ ,  $c_I$ , and  $c_M$  are three weights for controlling the importance of each item. Moreover,  $w_F(x)$ ,  $w_I(x)$ , and  $w_M(x)$  indicate the corresponding pixel-wise weight value for adjusting the contribution to loss function of each pixel.



**Fig. 4.** The examples of marker and interval mask extraction method. For visualizing, the value in  $D_e$  and  $D_s$  are normalized to a range of 0–255.

### 3.3. Mask extraction method for nuclei marker and interval

To achieve the marker annotation data, we firstly remove the intersection pixels of cell nuclei and transformed the labeled foreground (Fig. 4(a)) into binary image (Fig. 4(b)). Since the overlapping region of nuclei should be classified as the background when segmenting the foreground of nuclei. The binary foreground image would be used as the foreground ground truth. Then we compute the Euclidean distance map  $D_e$  (Fig. 4(c)) of the complement of binary foreground image, in which the value is smaller, the position is further away from the boundary. Therefore, we extract the regional minima of distance map  $D_e$  as the mask of marker  $G_M$  (Fig. 4(d)) by extended-minima transformation.

Based on the distance between two nuclei, we present a simple criterion to extract the interval mask (Fig. 4(f)) from the binary foreground image, as illustrated in [Algorithm 1](#). Firstly, we perform

---

#### Algorithm 1 The interval mask extraction algorithm.

**Input:** Binary foreground image:  $\{G_F(x), x \in \Omega\}$  including  $N$  cell nuclei

**Step 1.** Extracting the boundaries of nuclei  $B(x, N)$ .

**Step 2.** Computing the Euclidean distance between pixels and the boundaries of nuclei  $D(x, N)$ .

**Step 3.** Sorting  $D(x, N)$  and obtaining the distance of each pixel to the border of the second nearest nucleus  $D_s(x)$ .

**Step 4.** Computing the interval mask  $G_I(x)$  by the following criterion:

$$G_I(x) = \begin{cases} 1, & \text{if } \exp(-D_s(x)/\delta) \geq \alpha \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

**Step 5.**  $G_I(x) = G_I(x) \& \text{Not}\{G_F(x)\}$

**Output:** Interval ground truth  $G_I(x)$

---

morphological operation on the binary foreground image to obtain the boundary of each nucleus, and calculated the Euclidean distance between pixels and boundaries of nuclei. Then, we obtain the distance of each pixel to the second nearest boundary  $\{D_s(x), x \in \Omega\}$  (Fig. 4(e)) by sorting  $D(x, N)$ . The value of  $D_s(x)$  is smaller, meaning that the pixel located at  $x$  is closer to both two adjacent nuclei. We assume that the probability of pixel attribute to interval approximately meets the exponential distribution  $\exp(-D_s(x)/\delta)$ . Note that the distance value is greater, the probability value is

lower. Then, we utilize the threshold method with parameter  $\alpha$  to predict a pixel whether belong to interval. Finally, we perform the logical ‘and’ operation between the interval result and the inversion of foreground result to ensure that there are no foreground pixels incorrectly identified as the interval pixels.

To quantitatively analyze the influence of parameters  $\delta$  on the generation of interval mask, we show the change curve of  $\exp(-D_s(x)/\delta)$  with the pixel distance  $D_s$  in Fig. 5. The region limited by the blue dotted lines represents the main distribution range of well-accepted interval pixels. It is clear that setting  $\delta$  to 5 is conducive to extend the probability distribution range of interval pixels. In addition, we show the interval extraction results with setting the threshold  $\alpha$  to 0.3, 0.4, and 0.5 respectively. By comparing these results, we can find that the threshold method with small threshold value would misclassify some big gaps between the adjacent nuclei as the interval, while the big threshold value would lead to the increase of false negative results. According to the above findings, we set  $\alpha$  to 0.4 in our experiments.

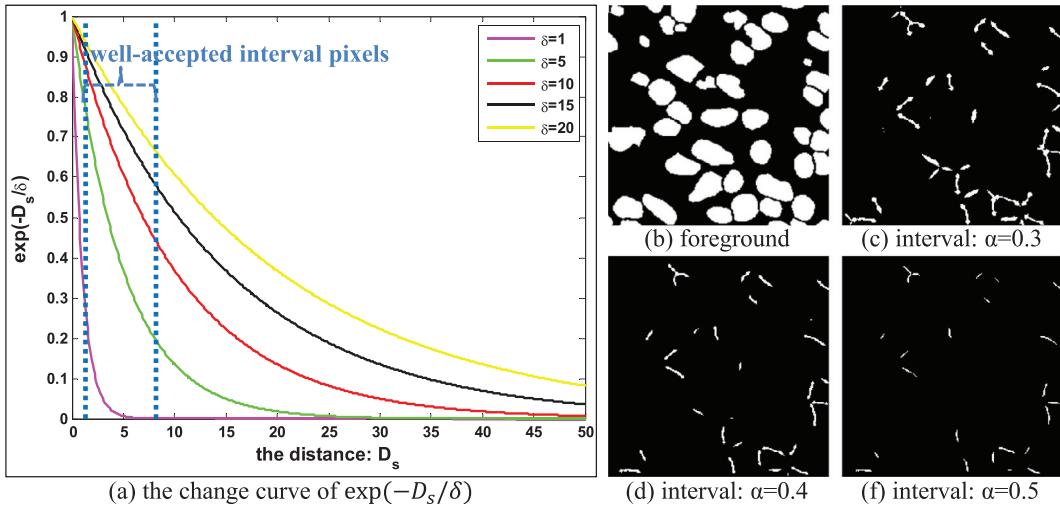
### 3.4. Pixel-wise weighting strategy

In practice, the data category distribution of most datasets are roughly inhomogeneous, called class-imbalance issue [35]. However, this issue would bias the learning algorithms toward the majority class, meaning that the algorithms tend to predict the minority class as the majority class. To reduce the negative effect caused by the imbalanced data, we apply the pixel-wise weighting strategies into our nuclei segmentation method. Firstly, we utilize the frequency of class to handle the imbalance problem between the foreground and non-foreground pixels. Assume that there are  $N_F$  foreground pixels in a nuclei image with total  $N_T$  pixels, we assigned the weight value of  $(N_T - N_F)/N_T$  and  $N_F/N_T$  to the foreground and background pixels. Moreover, we incorporate the distance map into the inverse class frequency method for generating the interval and marker weight maps as below:

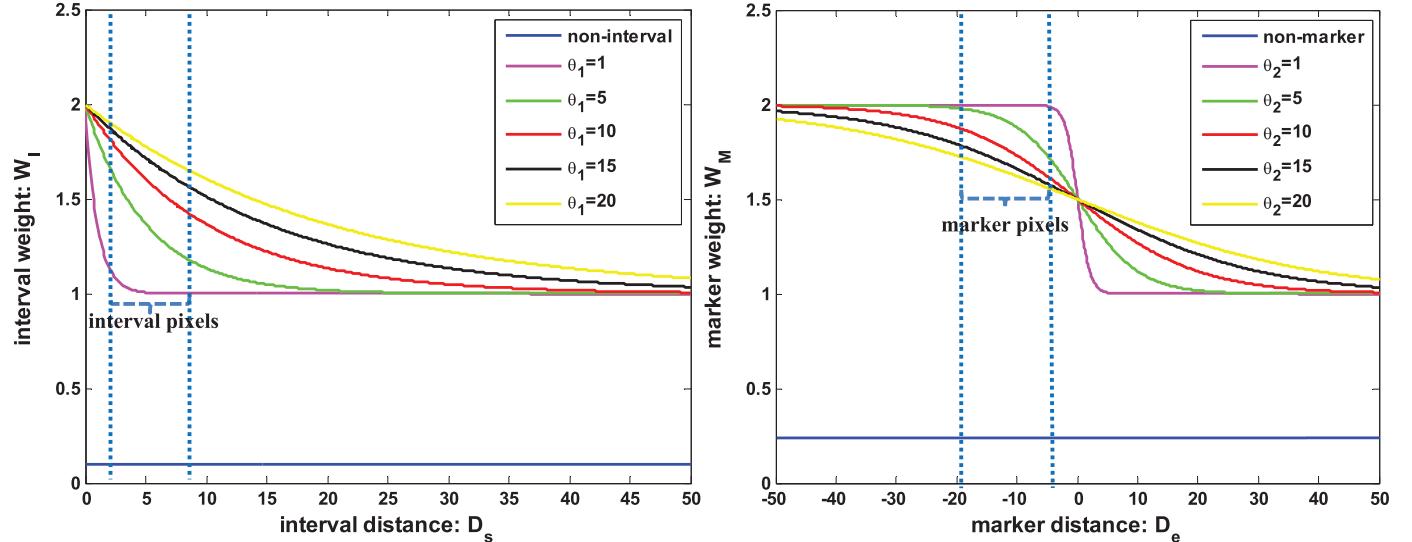
$$w_I(x) = \begin{cases} \frac{(N_F - N_I)}{N_T} + \exp(-D_s(x)/\theta_1), & \text{if } G_I(x) == 1 \\ \frac{N_I}{N_T}, & \text{otherwise} \end{cases} \quad (5)$$

$$w_M(x) = \begin{cases} \frac{(N_T - N_M)}{N_T} + \frac{1}{1 + \exp(D_e(x)/\theta_2)}, & \text{if } G_M(x) == 1 \\ \frac{N_M}{N_T}, & \text{otherwise} \end{cases} \quad (6)$$

where  $N_I$  and  $N_M$  denote the number of interval and marker pixels,  $\theta_1$  and  $\theta_2$  are two decay values for adjusting the contrast of weight



**Fig. 5.** The analysis of parameters in interval extraction method: (a) the change curve of  $\exp(-D_s(x)/\delta)$  with the pixel distance  $D_s$ , (b) the foreground of nuclei, and (c)–(f) the interval extraction results from (b) with setting three different threshold value.



**Fig. 6.** The varying curves of interval and marker weight value with the pixel distance value. The regions limited by the blue dotted lines represent the main distribution range of interval and marker pixels.

map. By this way, the closer to the center of nuclei, the weight value of marker is greater. And the value of interval weight map increases with the decline of the distance between two adjacent nuclei. These good properties can force our network to pay more attention to the smaller marker and interval objects, and thereby increase its accuracy of segmenting marker and interval regions.

To quantitatively analyze the influence of parameters  $\theta_1$  and  $\theta_2$  on the generation of marker and interval weight maps, we show the change curves of weight value accompanied with the distance value in Fig. 6. We can find that value of  $\theta_1$  and  $\theta_2$  directly determine the kurtosis and knee point of two curves. Due to that the interval pixels are mainly distributed around the distance value of 5, we set the parameter  $\theta_1$  to 5 for increasing the contrast of values in the weight map. For the same reason, we assign the parameter  $\theta_2$  to a value of 10.

### 3.5. Post-processing method

To further separate the overlapping nuclei in the nuclei foreground, we proposed an efficient post-processing method with fully taking advantages of the output results of DIMAN network and marker-controlled watershed method. The computing details

of this method are illustrated in Fig. 2. On one pathway, we perform a logical ‘and’ operation between the foreground result and the inversion of interval result for removing some overlapping regions of nuclei. Note that the logical operator is applied in a pixel-to-pixel way. Then the refined segmentation result would be converted into a distance map by Euclidean distance transformation method. On another pathway, the refined segmentation result is utilized to ensure that all the markers are in the foreground of nuclei by logical ‘and’ operation. Then each connected region in nuclei marker result is assigned with a unique integer value for representing different nucleus instances. Finally, we input the distance map, refined segmentation result, and labeled marker into the marker-controlled watershed algorithm for achieving the final nuclei segmentation result.

## 4. Implementation details

### 4.1. Datasets

In our experiments, we evaluate the performance of nuclei segmentation methods on three public histopathology image datasets:

**Dataset1<sup>1</sup>:** the training dataset of MICCAI 2017 Digital Pathology Challenge, which includes 32 annotated H&E stained histopathology image tiles of four different types of cancer. The nuclei images were extracted as rectangular tiles from the whole slide images at high resolution. Each labeled mask file is an array of integer values. The value of zero in a array represents the background, and the non-zero value means the corresponding pixel in the image tile is part of a nucleus. We randomly select 20 images from Dataset1 as the training data, and the rest as the testing data.

**Dataset2<sup>2</sup>:** presented by Naylor et al. [17], which contains 33 manually annotated H&E stained histopathology images with a total of 2754 cells. This dataset was obtained by randomly cropping the samples with size  $512 \times 512$  from the whole slide images. The overlapping nuclei in the image were annotated with different integer values by experts. The nuclei images in Dataset2 were randomly divided into two parts: training and validation data contains 23 samples, and the rest are the testing data.

**Dataset3<sup>3</sup>:** the multi-organ pathology images provided by Kumar et al. [32], which has been viewed as a benchmark for nuclei detection and segmentation. This dataset is consist of 30 samples with size  $1000 \times 1000$ , which were extracted from 30 whole slide images of 7 different organs. The training and validation set in Dataset3 contains 16 images of 4 kinds of organs, and the testing set contains 14 images of 7 kinds of organs. This partition method can examine well the robustness and generalization ability of models.

#### 4.2. Data augmentation methods

Compared with the nature image classification and segmentation tasks, the number of training data in our experiments is very small. However, training a deep convolutional neural network needs a large number of image data in avoid of over-fitting. In order to solve this problem, some data augmentation methods were used to expand the training data in our experiments. Firstly, we extracted numerous image patches with size  $224 \times 224$  from the original images by random cropping method. In order to ensure that each image patch has positive samples (foreground pixels), the center of image patch is located at the foreground. Then we performed the vertical and horizontal flipping on the patches with 50% probability, which can enhance the robustness of proposed network. In addition, we also adopted the removing mean value and min-max normalization operations to pre-process the training data for reducing the diversity of pixel intensities in nuclei images.

#### 4.3. Training and inference

We implemented our method based on the open-source machine learning framework TensorFlow [36] and deep learning library TensorLayer [37]. All experiments were conducted on a graphic workstation with 3.5GHz Intel Xeon CPU and NVIDIA M4000 GPU. During the training process, we selected the Adam optimization algorithm to minimize the loss function of our network. The hyper-parameters related with training stage were set as follows: batch size of training data (12), learning rate (initiated as 0.0002 and decreased by a factor 0.8 until 0.000001), dropout rate (0.6), and the exponential decay rate for the first and second moment estimates (0.8 and 0.999). Specially, we assigned the weight vector  $[c_F, c_I, c_M]$  in the loss function of DIMAN network to a value of [1,1,1]. Training DIMAN network 300 epochs using about 500 samples costs 4.5 hours. The training progress of DIMAN net-

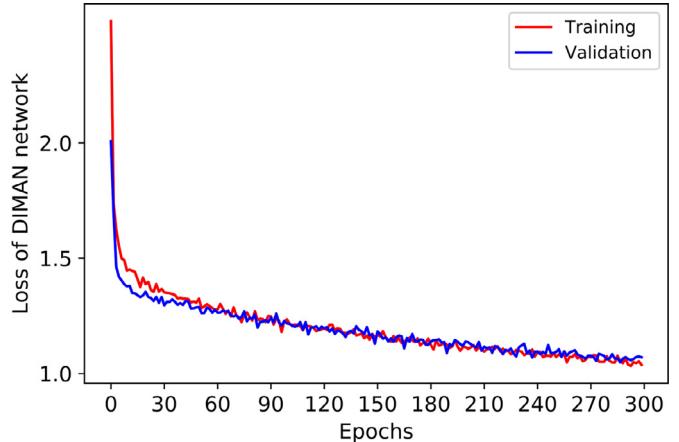


Fig. 7. The training and validation loss curves of DIMAN network..

work is illustrated in Fig. 7. We can find that our network has a fast convergence speed, good stability and generalization ability.

At inference stage, the testing image need to be resized to  $512 \times 512$  and to be pre-processed by the removing mean value and min-max normalization methods. Then our network outputs the posteriori probability of multiple segmentation results by forward propagation. We transformed the probability map into binary segmentation image by thresholding method with setting the threshold value to 0.5. Finally, the proposed post-processing method was used to process the multiple segmentation results of DIMAN network for achieving the final segmentation result, in which each connection region with a unique integer value represents a cell nucleus instance. The average processing time per image with size  $512 \times 512$  is 0.72 s for our method. Based on the proposed model, we have built a real-time processing tool for segmenting the overlapping nuclei in the region-of-interest of whole slide images.

#### 4.4. Evaluation metrics

We applied multiple widely regarded metrics to quantitatively analyze the nuclei segmentation performance on Dataset1 and Dataset2 for all methods, including F1-score ( $F_1$ ), Precision rate ( $P$ ), Recall rate ( $R$ ), Object-level dice index ( $ODI$ ), and Object-level hausdorff distance ( $OHD$ ). Note that these evaluation metrics are implemented by matlab programming language and provided by the gland segmentation challenge contest<sup>4</sup>. In addition, the Aggregated jaccard index (AJI) [32] was employed to access the nuclei segmentation performance on Dataset3 for comparing with other state-of-the-arts methods. This metric is a modification of the global Jaccard index, which can evaluate the pixel-level and object-level nuclei segmentation performance in a unified framework.

### 5. Results and analysis

#### 5.1. Comparison with other networks and post-processing methods

We compare our proposed approach with four known CNN-based semantic segmentation methods, including FCN8s [38], HED [39], Unet [40], and SharpMask [41]. For fairness, all competing networks were trained under the same conditions with our method. In addition, the conditional erosion based watershed (CEW [28]) method [28], morphological dynamics based watershed

<sup>1</sup> available at <http://miccai.cloudapp.net/competitions/>.

<sup>2</sup> available at <https://peterjacknaylor.github.io/>.

<sup>3</sup> available at <https://nucleisegmentationbenchmark.weebly.com/>.

<sup>4</sup> <https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/evaluation/>.

**Table 1**

Comparison of Nuclei Detection and Segmentation on Dataset1. All values are expressed as “mean value/standard deviation”.

Networks	Post-processing	Nuclei Detection			Nuclei Segmentation	
		F <sub>1</sub>	P	R	ODI	OHD
FCN8s [38]	–	0.74/0.078	0.749/0.089	0.747/0.131	0.693/0.093	16.196/5.72
	FES [42]	0.604/0.188	0.921/0.048	0.474/0.197	0.589/0.12	19.411/7.549
	CPS [43]	0.754/0.107	0.722/0.118	0.793/0.107	0.704/0.063	10.296/3.376
	CEW [28]	0.761/0.05	0.69/0.08	0.859/0.079	0.729/0.054	12.018/2.326
	MDW [17]	0.785/0.068	0.712/0.105	0.889/0.064	0.754/0.057	9.954/1.357
HED [39]	–	0.781/0.114	0.843/0.055	0.739/0.166	0.7/0.128	17.535/9.634
	FES [42]	0.639/0.155	0.915/0.071	0.515/0.177	0.599/0.12	20.123/7.215
	CPS [43]	0.817/0.098	0.805/0.131	0.839/0.085	0.752/0.056	10.506/5.755
	CEW [28]	0.813/0.081	0.791/0.055	0.841/0.123	0.745/0.096	12.663/6.113
	MDW [17]	0.844/0.056	0.806/0.06	0.887/0.068	0.78/0.061	9.804/2.84
Unet [40]	–	0.74/0.092	0.732/0.086	0.763/0.15	0.7/0.108	16.343/8.056
	FES [42]	0.648/0.172	<b>0.948/0.053</b>	0.515/0.179	0.605/0.105	19.276/6.312
	CPS [43]	0.804/0.067	0.788/0.085	0.826/0.086	0.732/0.051	10.957/4.392
	CEW [28]	0.77/0.066	0.693/0.088	0.878/0.095	0.74/0.066	11.889/4.367
	MDW [17]	0.736/0.088	0.702/0.098	0.788/0.138	0.712/0.099	14.977/7.243
SharpMask [41]	–	0.702/0.102	0.676/0.148	0.765/0.148	0.699/0.107	16.546/6.862
	FES [42]	0.624/0.156	0.933/0.053	0.487/0.168	0.577/0.107	21.025/5.944
	CPS [43]	0.768/0.112	0.73/0.139	0.822/0.099	0.707/0.06	11.666/5.289
	CEW [28]	0.723/0.103	0.641/0.142	0.858/0.105	0.734/0.077	12.857/4.403
	MDW [17]	0.678/0.133	0.582/0.179	0.862/0.103	0.742/0.08	12.026/4.065
DIMAN	–	0.705/0.101	0.676/0.128	0.762/0.145	0.695/0.106	18.875/10.014
	Removing Interval	0.732/0.086	0.687/0.122	0.806/0.115	0.723/0.083	15.01/5.179
	FES [42]	0.628/0.155	0.9/0.071	0.51/0.184	0.598/0.115	20.231/6.879
	CPS [43]	0.805/0.073	0.767/0.098	0.857/0.081	0.75/0.045	10.087/3.689
	CEW [28]	0.747/0.078	0.644/0.113	0.907/0.07	0.757/0.053	11.384/3.098
	MDW [17]	0.714/0.111	0.604/0.151	<b>0.907/0.057</b>	0.769/0.049	10.439/2.195
	Proposed	<b>0.886/0.032</b>	0.881/0.039	0.892/0.041	<b>0.781/0.05</b>	<b>9.467/1.622</b>

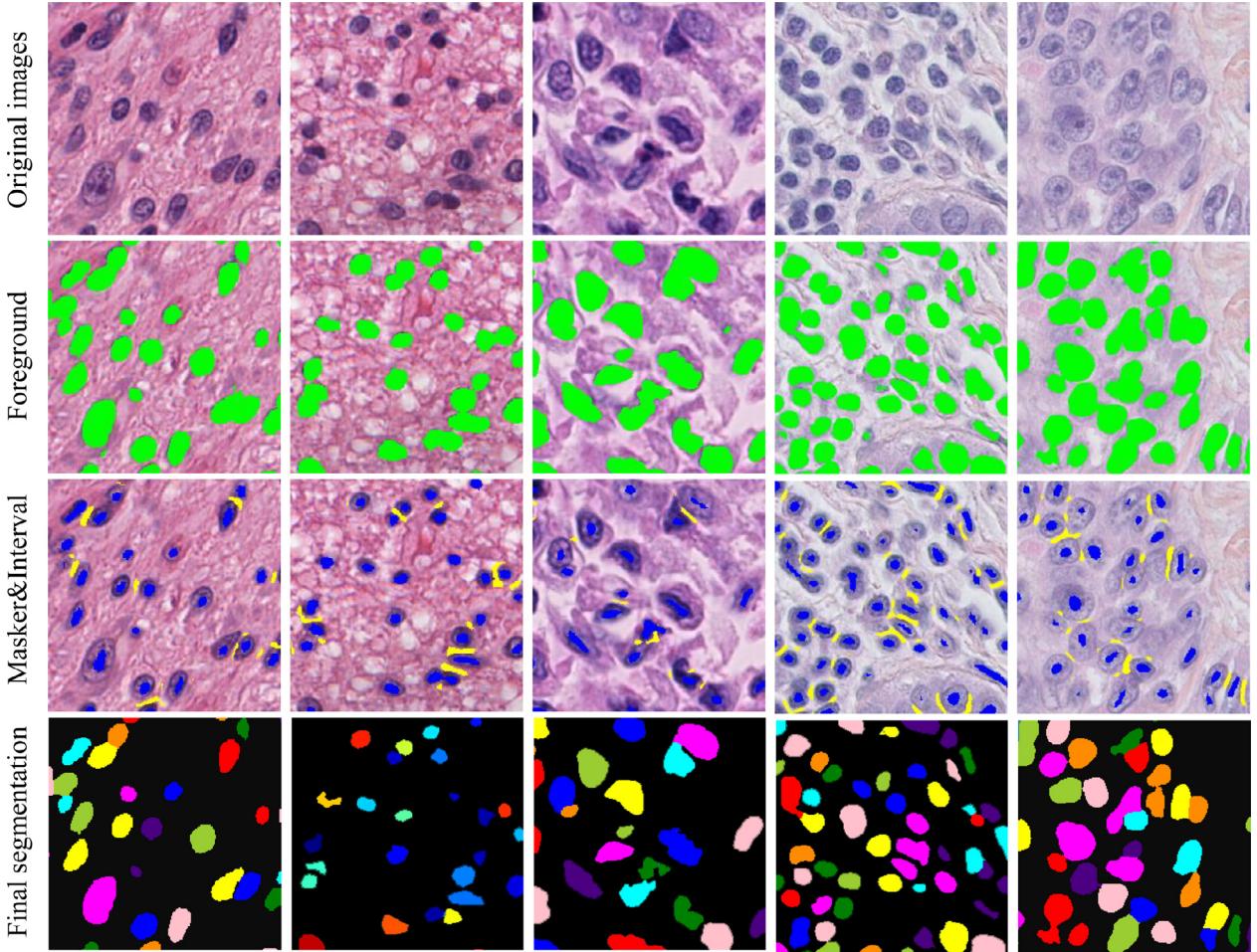
**Table 2**

Comparison of Nuclei Detection and Segmentation on Dataset2. All values are expressed as “mean value/standard deviation”.

Networks	Post-processing	Nuclei Detection			Nuclei Segmentation	
		F <sub>1</sub>	P	R	ODI	OHD
FCN8s [38]	–	0.733/0.091	0.733/0.087	0.742/0.124	0.694/0.072	14.746/6.292
	FES [42]	0.624/0.212	0.915/0.058	0.508/0.246	0.603/0.175	19.499/9.373
	CPS [43]	0.795/0.105	0.781/0.122	0.819/0.124	0.72/0.09	11.065/2.988
	CEW [28]	0.742/0.091	0.681/0.109	0.824/0.094	0.722/0.059	11.79/5.839
	MDW [17]	0.751/0.073	0.682/0.089	0.845/0.089	0.74/0.046	9.907/3.178
HED [39]	–	0.765/0.099	0.832/0.074	0.726/0.157	0.709/0.095	17.135/9.631
	FES [42]	0.637/0.218	0.937/0.033	0.52/0.252	0.596/0.192	21.025/10.319
	CPS [43]	0.852/0.067	0.853/0.057	0.853/0.088	0.759/0.088	10.122/3.759
	CEW [28]	0.805/0.062	0.794/0.082	0.828/0.106	0.763/0.046	10.986/4.548
	MDW [17]	0.815/0.066	0.781/0.094	0.862/0.076	0.78/0.044	10.422/4.993
Unet [40]	–	0.704/0.092	0.732/0.114	0.711/0.163	0.69/0.104	18.842/9.949
	FES [42]	0.621/0.205	<b>0.941/0.046</b>	0.505/0.25	0.589/0.178	20.583/9.042
	CPS [43]	0.824/0.066	0.793/0.058	0.86/0.09	0.74/0.077	10.395/2.968
	CEW [28]	0.739/0.104	0.701/0.15	0.809/0.105	0.731/0.062	14.308/6.395
	MDW [17]	0.708/0.085	0.72/0.116	0.725/0.147	0.699/0.094	17.773/8.831
SharpMask [41]	–	0.676/0.097	0.665/0.127	0.722/0.153	0.679/0.099	18.958/10.362
	FES [42]	0.631/0.185	0.907/0.054	0.516/0.229	0.601/0.177	19.94/8.785
	CPS [43]	0.807/0.069	0.776/0.088	0.849/0.094	0.744/0.092	10.399/3.573
	CEW [28]	0.7/0.117	0.636/0.157	0.816/0.116	0.728/0.061	13.188/6.046
	MDW [17]	0.644/0.141	0.557/0.163	0.806/0.12	0.722/0.07	13.204/6.15
DIMAN	–	0.735/0.091	0.73/0.115	0.765/0.14	0.72/0.084	16.752/8.581
	Removing Interval	0.751/0.087	0.734/0.117	0.79/0.119	0.739/0.066	14.795/6.345
	FES [42]	0.625/0.206	0.923/0.05	0.512/0.248	0.591/0.185	20.552/9.245
	CPS [43]	0.833/0.057	0.791/0.074	<b>0.883/0.065</b>	0.765/0.072	10.103/3.101
	CEW [28]	0.769/0.105	0.701/0.141	0.87/0.078	0.778/0.04	10.326/3.156
	MDW [17]	0.728/0.098	0.643/0.129	0.86/0.074	0.775/0.037	10.267/3.177
	Proposed	<b>0.87/0.075</b>	0.87/0.1	0.875/0.069	<b>0.785/0.035</b>	<b>9.886/4.763</b>

(MDW [17]) method [17], fitting ellipses based segmentation (FES) method [42], and concave points based segmentation (CPS) method [43] were used as the post-processing methods for refining the nuclei segmentation results. In Tables 1 and 2, we show the nuclei detection and segmentation performance of all methods on two public histopathology image datasets.

The experimental results prove that the DIMAN network with proposed post-processing method outperforms other competing methods on both two datasets in terms of nuclei detection and segmentation. Especially, the F1-score of our method on Dataset1 and Dataset2 are 0.886/0.032 and 0.87/0.075 with achieving +4.98% and +2.11% gain than the second one. Moreover, we



**Fig. 8.** Intermediate and final segmentation results of sub-images of our method. The third row is the combination of interval (yellow) and marker (blue) segmentation results.

can observe that the proposed approach has obvious advantage in terms of Precision and Recall rate, which would help to increase the accuracy of the nuclei counting. In addition, the highest Object-level dice index value and the lowest Object-level hausdorff distance value of our method demonstrate that the contours of our segmentation result are more accordant with the manual delineated result than other approaches. This advantage is conducive to the shape information extraction and morphological measurement of nuclei.

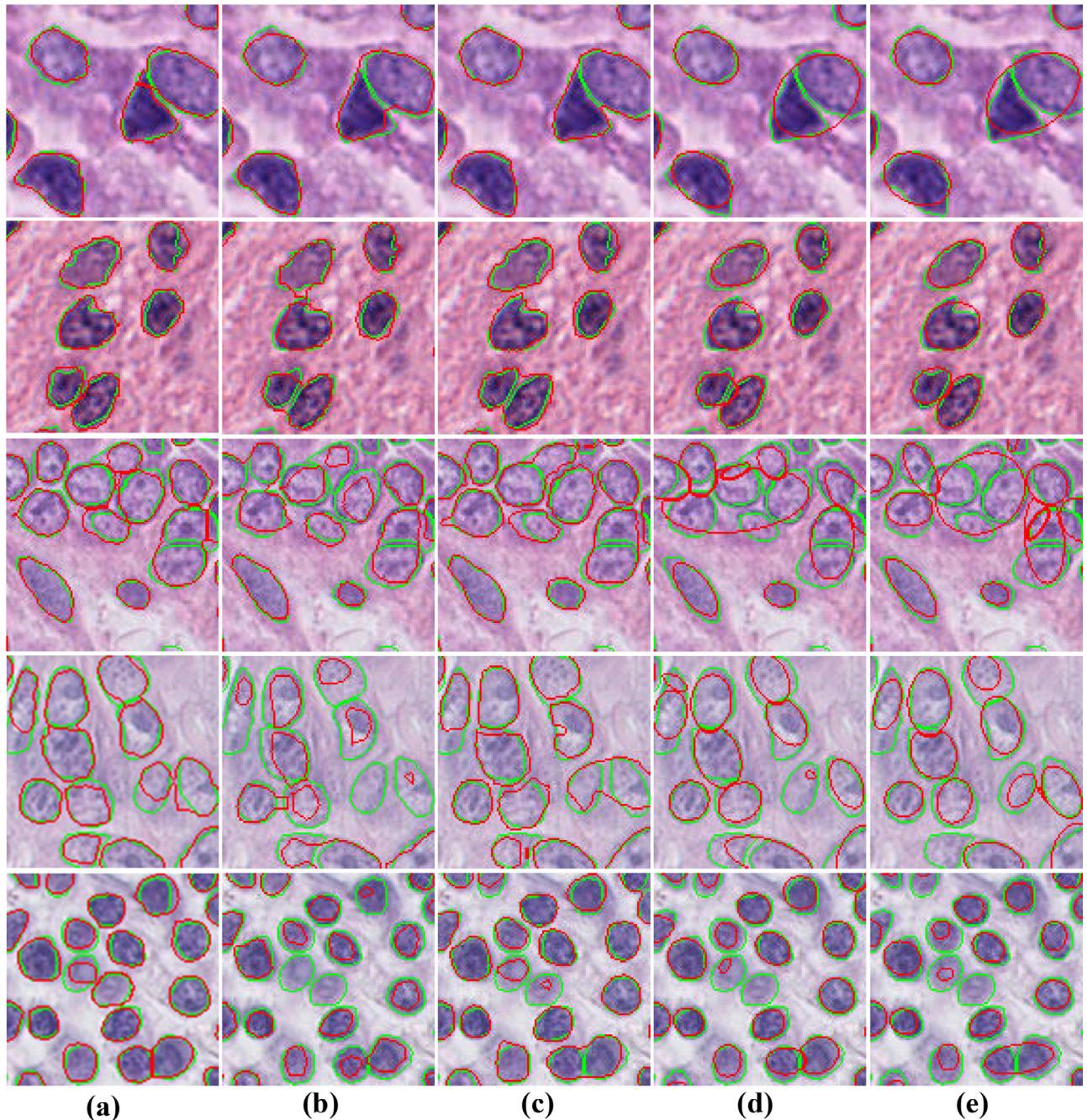
The comparison between DIMAN and DIMAN with removing interval operation shows that the interval information between the adjacent nuclei can efficiently improve the nuclei detection and segmentation performance especially in the term of Recall rate and Object-level hausdorff distance. The improvement of Recall rate means that some overlapping nuclei in the foreground were separated successfully. Moreover, by comparing the evaluation values of DIMAN network with different post-processing methods, we can find that the proposed post-processing method has the best effect in improving the nuclei segmentation performance. The main reason is that the nuclei marker result of proposed approach is more reasonable and accurate than the other post-processing methods, which extract the nuclei marks from the foreground segmentation result rather than the original image.

To illustrate the nuclei segmentation effect of our method in details, we shows the intermediate results and final segmentation results of five representative image patches of our method in Fig. 8. We can clearly see that: (1) the foreground segmentation results

(green) basically cover the region of nuclei; (2) most of the interval segmentation results locate between two adjacent nuclei and have the ability to sperate the overlapping nuclei, (3) in the center of each nucleus instance there is a segmented marker object; (4) the majority of crowded nuclei are successfully divided into multiple individual nucleus objects by our method, which is in good agreement with ground truth. Moreover, we show five representative nuclei segmentation examples in Fig. 9 for intuitively comparing the nuclei segmentation performance of four other methods with the proposed method. It is clear that our method can efficiently segment the touching nuclei, and has better coherence with the ground truth than other approaches. By contrast, the under-segmentation problem is obvious in the segmentation results of other methods especially for the SharpMask with CEW method.

### 5.2. Comparison with the state-of-arts methods

To compare our method with five state-of-the-arts nuclei segmentation methods, we utilized our method to segment the testing images on Dataset3 and computed the Aggregated jaccard index value of each nuclei segmentation result, as illustrated in Table 3. For the state-of-the-arts methods, we directly cite the AJI values recorded in the original literatures. Especially, the AJI values of CNN2 and CNN3 methods have been viewed as the benchmark [32]. The DIST algorithm proposed by Naylor et al. [20] achieved a better performance in nuclei segmentation than CNN2 and CNN3. The Unet [40] and Mask R-CNN [44] networks were utilized as the



**Fig. 9.** Comparative nuclei segmentation results of various methods: (a) DIMAN with proposed post-processing method, (b) FCN8s with MDW method, (c) HED with MDW method, (d) Unet with CPS method, and (e) SharpMask with CPS method. The green and red contours represent the ground truth and segmentation result respectively.

comparison methods in their experiments. The average AJI value of CNN3 method is higher than Mask R-CNN network but lower than Unet network.

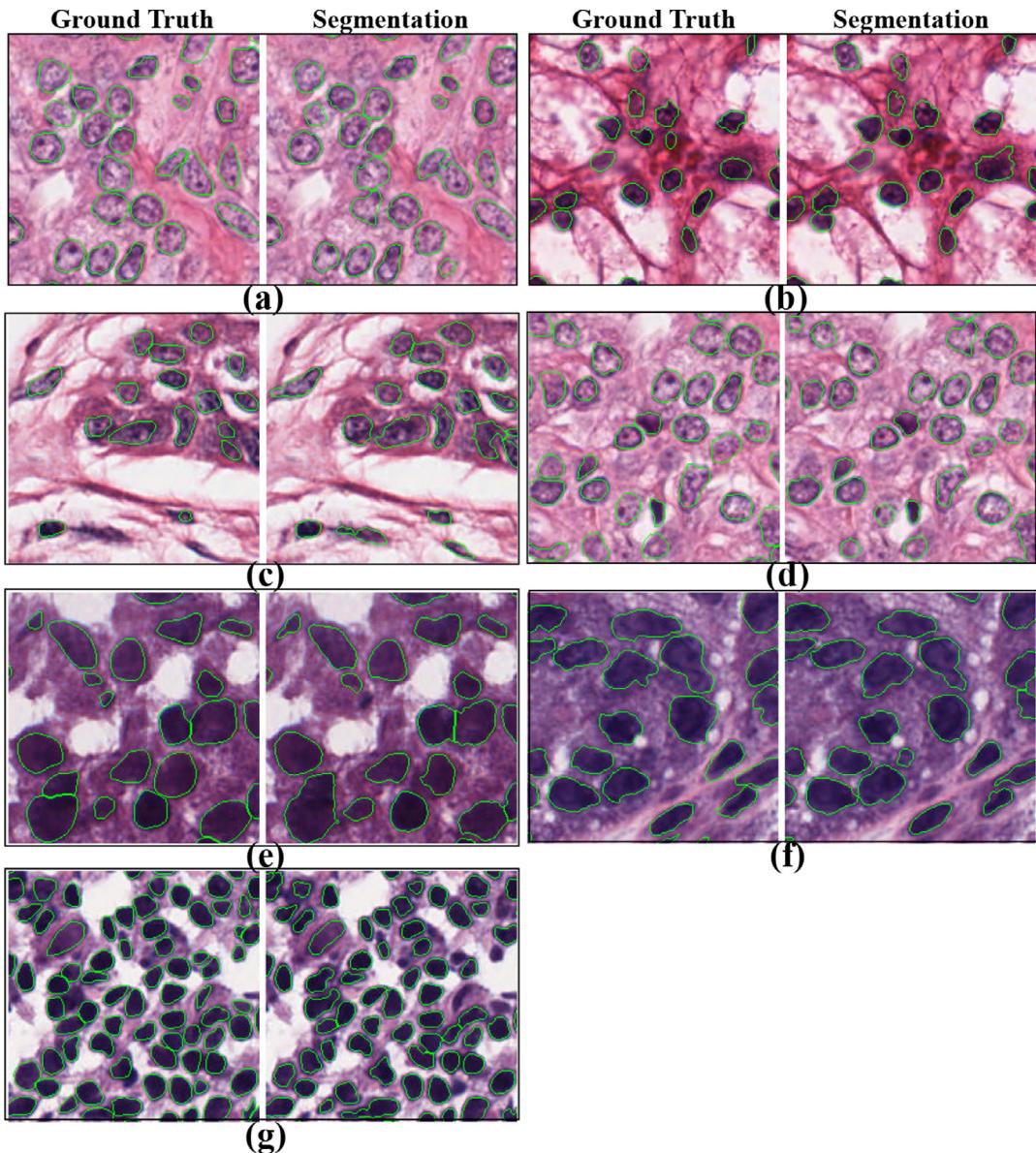
According to the experimental data, we can find that the mean value of aggregated jaccard index of our method is higher than other methods, and our approach has obvious predominance on segmenting the histopathology images of liver and prostate. The proposed approach achieves a +62.67% improvement compared with CNN2, +11.43% improvement compared with CNN3, and +9.3% improvement compared with Unet. Moreover, our approach obtains a bit of progress in nuclei segmentation as compared to the state-of-the-art DIST method, and the standard deviation of our method is lower than the DIST, which indicates that the nuclei segmentation performance of our approach is more stable than

DIST method. Furthermore, we show the nuclei seven representative segmentation results of our method, which are from different organs in Fig. 10. It is clear that the boundaries of segmentation results have outstanding agreement with the ground truth. The above experimental data can prove the accuracy and reliability of our approach in overlapping nuclei segmentation.

### 5.3. Analysis of experimental details

#### 5.3.1. Setting of parameters in loss function

In order to analyze the relationship between the segmentation performance of proposed method and the weight vector  $[c_F, c_I, c_M]$  in the loss function, we performed a series of comparative experiments on Dataset1. As illustrated in Fig. 11, we learned multi-

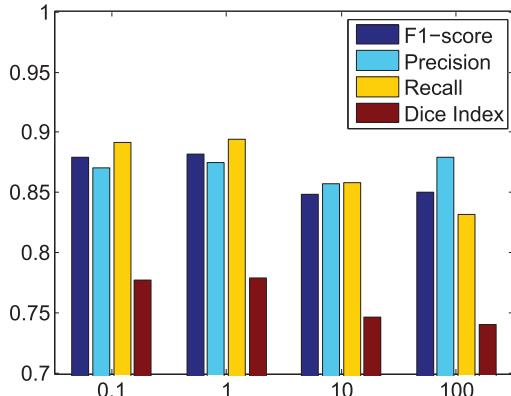
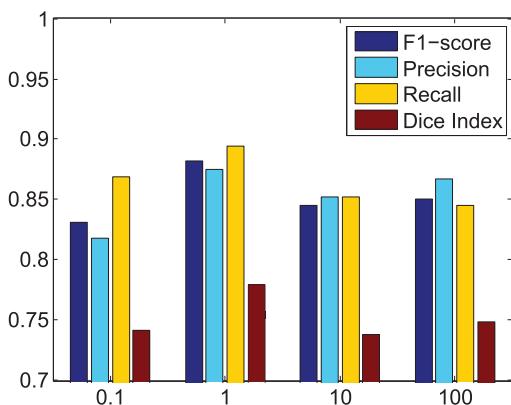
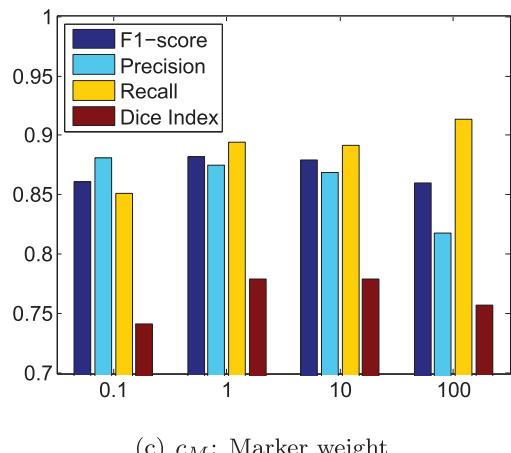


**Fig. 10.** Qualitative Results on Dataset3. The histopathology images are from seven different organs: (a) breast, (b) liver, (d) kidney, (d)prostate, (e) bladder, (f) colon, (g) stomach. The boundaries of nuclei are annotated with green curves.

**Table 3**

The Aggregated jaccard index comparison of our method with five state-of-the-arts methods on Dataset3.

Organ	Image	Methods					
		CNN2 [32]	CNN3 [32]	Unet [20]	Mask R-CNN [20]	DIST [20]	Proposed
Breast	1	0.3852	0.4974	0.5510	0.4515	<b>0.6475</b>	0.4999
	2	0.4663	<b>0.5796</b>	0.5296	0.5311	0.5467	0.5523
Liver	1	0.4086	0.5175	0.3751	0.4861	0.4240	<b>0.6334</b>
	2	0.3325	0.5148	0.4371	0.4383	0.4484	<b>0.5633</b>
Kidney	1	0.3129	0.4792	<b>0.6515</b>	0.5600	0.6408	0.4719
	2	0.5010	<b>0.6672</b>	0.6543	0.4764	0.6550	0.5357
Prostate	1	0.2707	0.4914	0.4559	0.4970	0.5334	<b>0.5740</b>
	2	0.1848	0.3761	0.4803	0.5673	0.5884	<b>0.6208</b>
Bladder	1	0.3498	0.5465	0.5815	0.5693	0.5648	<b>0.6738</b>
	2	0.2876	0.4968	0.5037	0.4328	<b>0.5420</b>	0.5337
Colon	1	0.3043	0.4891	0.4846	0.3624	<b>0.5466</b>	0.5167
	2	0.3125	<b>0.5692</b>	0.3721	0.4003	0.4432	0.5430
Stomach	1	0.3961	0.4538	0.5831	0.5983	<b>0.6273</b>	0.5957
	2	0.3618	0.4378	0.5945	<b>0.6318</b>	0.6294	0.6150
Overall (mean/std)		0.3482/0.0813	0.5083/0.0695	0.5182/0.0906	0.5002/0.0791	0.5598/0.0781	<b>0.5664/0.0561</b>

(a)  $c_F$ : Foreground weight(b)  $c_I$ : Interval weight(c)  $c_M$ : Marker weight

**Fig. 11.** The performance of the proposed segmentation approach with changing the foreground weight (a), interval weight (b) and marker weight (c) from 0.1 to 100.

ple networks with different weight vectors by varying one weight value from 0.1 to 100 with keeping other two weights equal to 1. The experimental data shows that the choice of weight vector has an obvious impact on the performance of nuclei detection and segmentation. For instance, the F1-score and Object-level dice index of our method obtain a great growth, when the weight of  $c_I$  increases from 0.1 to 1. Since the increasing of  $c_I$  can force the network to pay more attention to interval segmentation task, which has positive effect on the overlapping nuclei segmentation. However, when

**Table 4**  
Comparison experiments for weighting strategies.

Method	Weighting strategy	$F_1$	$P$	$R$	$ODI$	$OHD$
DIMAN+MCW	focal loss [45]	0.664	0.697	0.634	0.658	20.872
DIMAN+MCW	inverse class frequency	0.808	0.822	0.795	0.725	12.239
DIMAN+MCW	proposed	0.878	0.875	0.882	0.77	10.276

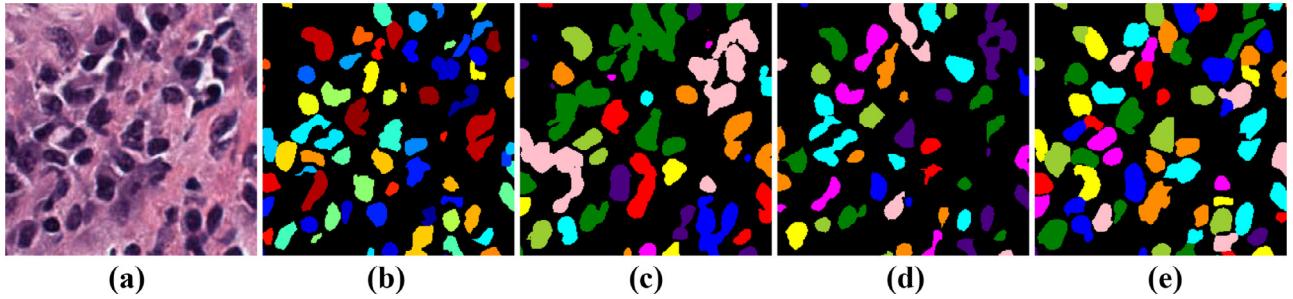
one weight is far greater than others, the evaluation scores show a clear downward trend. Another important phenomenon is that our approach achieves the best performance when the weight vector is set to [1,1,1]. Overall, the above experimental data suggests that the final segmentation performance of our method should be decided together by the foreground, interval, and marker segmentation tasks, and the roles of all items in the loss function should be equal for achieving the best result.

### 5.3.2. Comparison with other weighting strategies

To prove the effectiveness of our weighting strategy, we compared our method with other two widely used weighting strategies, including the inverse class frequency (ICF) and focal loss (FL) [45] methods. In the inverse class frequency weighting strategy, the misclassified loss value would be multiplied by the inverse of class frequency. By contrast, the focal loss strategy is to dynamically decrease the weight value of well-classified examples by incorporating a factor  $(1 - p(x))^\gamma$  into the loss function. We applied the two weighting strategies into the loss function of DIMAN network. In Table 4, we show the metric values evaluated on the Dataset1 for three different weighting strategies. The experimental data reveals that the DIMAN network with the proposed weighting strategy has better performance in nuclei segmentation than the DIMAN network with inverse class frequency and focal loss weighting strategies. Especially, our method obtains +8.66% and +6.21% increase in terms of F1-score and Object-level dice index compared with the inverse class frequency method. However, the DIMAN network with focal loss weighting strategy obtains the worst performance in nuclei segmentation. In addition, we show the comparison results for three different weighting strategies in Fig. 12. Obviously, the nuclei segmentation result of our method achieves the better agreement with the ground truth.

### 5.3.3. Impact of data augmentation methods

To analyze the influence of data augmentation methods on nuclei segmentation, we designed some comparative experiments on Dataset2 as illustrated in Table 5. In random cropping approach, the parameter “crop number” indicates the number of cropping patches from each original image. The total number of image patches (denoted by “total number”) sampled from the training data is proportional to this parameter. Note that the random vertical and horizontal flipping methods can increase the diversity of data but have no effect on the number of total patches. We applied these data augmentation methods into the training stage of DIMAN network, and calculated the F1-scores and Object-level dice index values of our method for comparison. The big gap between the evaluation values in the first and last row shows that the data augmentation methods can efficiently promote the nuclei segmentation performance for our method. Especially, the DIMAN method with three data augmentation methods achieves 0.865 F1-score, which is 0.213 higher than the DIMAN method without the usage of data augmentation methods. By comparing the data from the second to fourth row in Table 5, we see that the evaluation values are increasing with the number of cropping patches. However, the gain generated by random cropping approach decreases along with



**Fig. 12.** Comparison results for weighting strategies: (a) original image, (b) ground truth, (c) focal loss, (d) inverse class frequency, (e) proposed.

**Table 5**

Comparison experiments for data augmentation methods. Note that '✓' represents that the corresponding augmentation method was used, while the meaning of '✗' is opposite.

Random crop	Vertical flip	Horizontal flip	Crop number	Total number	$F_1$	ODI
✗	✗	✗	✗	20	0.652	0.625
✓	✗	✗	10	200	0.761	0.693
✓	✗	✗	20	400	0.814	0.728
✓	✗	✗	30	600	0.826	0.742
✓	✓	✗	20	400	0.846	0.75
✓	✓	✓	20	400	0.865	0.762

the increase of total number of patches. In addition, the improvement effect of vertical and horizontal flipping in nuclei segmentation is lower than the random cropping method.

## 6. Conclusion and future work

It's well known that overlapping nuclei is one of the major challenges for nuclei segmentation in histopathology images. For tackling this problem, we have established a novel computational framework by combining the convolutional neural networks with marker-controlled watershed. Unlike other competing methods, which extract the markers of nuclei from the foreground segmentation result, we proposed a deep CNN architecture for predicting the foreground, interval and marker of nuclei simultaneously. There are at least two advantages in our proposed method: (1) no need for a special marker extraction method, which simplifies the flowchart and improves the efficiency, and (2) obtaining better marker results, which can boost the performance of MCW. The experimental results prove that our approach outperforms other state-of-art methods in nuclei detection and segmentation. However, the shape of nuclei in segmentation result is irregular, which is detrimental to subsequent morphological analysis of nuclei. In future work, we intend to integrate the deep convolutional neural networks with other advanced segmentation methods for segmenting the overlapping nuclei, and take the related domain knowledge of medicine into our model for improving the robustness of model.

## Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work is supported by NSFC (Grant numbers G0561671135, G0591630311, M050102011531005). The authors thank the anonymous reviewers for their valuable comments.

## References

- [1] M. Veta, J.P.W. Pluim, P.J. van Diest, M.A. Viergever, Breast cancer histopathology image analysis: a review, *IEEE Trans. Biomed. Eng.* 61 (5) (2014) 1400–1411, doi:[10.1109/tbme.2014.2303852](https://doi.org/10.1109/tbme.2014.2303852).
- [2] R.G. Vincent, J.W. Pickren, W.W. Lane, I. Bross, H. Takita, L. Houten, A.C. Gutierrez, T. Rzepka, The changing histopathology of lung cancer. a review of 1682 cases, *Cancer* 39 (4) (1977) 1647–1655, doi:[10.1002/1097-0142\(197704\)39:4<1647::AID-CNCR2820390439>3.0.CO;2-H](https://doi.org/10.1002/1097-0142(197704)39:4<1647::AID-CNCR2820390439>3.0.CO;2-H).
- [3] S. Naik, S. Doyle, S. Agner, A. Madabhushi, M. Feldman, J. Tomaszewski, Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology, in: Proceedings of the 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, 2008, doi:[10.1109/isbi.2008.4540988](https://doi.org/10.1109/isbi.2008.4540988).
- [4] H.A. Alturkistani, F.M. Tashkandi, Z.M. Mohammedsaleh, Histological stains: a literature review and case study, *Glob. J. Health Sci.* 8 (3) (2015) 72, doi:[10.5539/gjhs.v8n3p72](https://doi.org/10.5539/gjhs.v8n3p72).
- [5] M.N. Gurcan, L.E. Boucheron, A. Can, A. Madabhushi, N.M. Rajpoot, B. Yener, Histopathological image analysis: a review, *IEEE Rev. Biomed. Eng.* 2 (2009) 147–171, doi:[10.1109/RBME.2009.2034865](https://doi.org/10.1109/RBME.2009.2034865).
- [6] M. Veta, P.J. Van Diest, S.M. Willems, H. Wang, A. Madabhushi, A. Cruz-Roa, F. Gonzalez, A.B. Larsen, J.S. Vestergaard, A.B. Dahl, et al., Assessment of algorithms for mitosis detection in breast cancer histopathology images, *Med. Image Anal.* 20 (1) (2015) 237–248, doi:[10.1016/j.media.2014.11.010](https://doi.org/10.1016/j.media.2014.11.010).
- [7] L. Pantanowitz, P.N. Valenstein, A.J. Evans, K.J. Kaplan, J.D. Pfeifer, D.C. Wilbur, L.C. Collins, T.J. Colgan, Review of the current state of whole slide imaging in pathology, *J. Pathol. Inf.* 2 (2011), doi:[10.4103/2153-3539.83746](https://doi.org/10.4103/2153-3539.83746).
- [8] V.-T. Ta, O. Lézoray, A. Elmoataz, S. Schüpp, Graph-based tools for microscopic cellular image segmentation, *Pattern Recognit.* 42 (6) (2009) 1113–1125, doi:[10.1016/j.patcog.2008.10.029](https://doi.org/10.1016/j.patcog.2008.10.029).
- [9] O. Sertel, J. Kong, H. Shimada, U. Catalyurek, J.H. Saltz, M. Gurcan, Computer-aided prognosis of neuroblastoma: classification of stromal development on whole-slide images, in: M.L. Giger, N. Karssemeijer (Eds.), *Medical Imaging 2008: Computer-Aided Diagnosis*, 6915, SPIE, 2008 69150P, doi:[10.1117/12.770666](https://doi.org/10.1117/12.770666).
- [10] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, R. Monczak, Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images, *Comput. Biol. Med.* 43 (10) (2013) 1563–1572, doi:[10.1016/j.combiomed.2013.08.003](https://doi.org/10.1016/j.combiomed.2013.08.003).
- [11] F. Xing, Y. Xie, L. Yang, An automatic learning-based framework for robust nucleus segmentation, *IEEE Trans. Med. Imaging* 35 (2) (2016) 550–566, doi:[10.1109/tmi.2015.2481436](https://doi.org/10.1109/tmi.2015.2481436).
- [12] X. Qi, F. Xing, D.J. Foran, L. Yang, Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set, *IEEE Trans. Biomed. Eng.* 59 (3) (2012) 754–765, doi:[10.1109/tbme.2011.2179298](https://doi.org/10.1109/tbme.2011.2179298).
- [13] Y. Al-Kofahi, W. Lassoued, W. Lee, B. Roysam, Improved automatic detection and segmentation of cell nuclei in histopathology images, *IEEE Trans. Biomed. Eng.* 57 (4) (2010) 841–852, doi:[10.1109/tbme.2009.2035102](https://doi.org/10.1109/tbme.2009.2035102).
- [14] J. Xu, L. Xiang, Q. Liu, H. Gilmore, J. Wu, J. Tang, A. Madabhushi, Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology im-

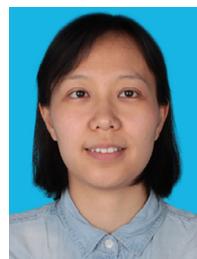
- ages, *IEEE Trans. Med. Imaging* 35 (1) (2016) 119–130, doi:[10.1109/tmi.2015.2458702](https://doi.org/10.1109/tmi.2015.2458702).
- [15] A. Janowczyk, A. Madabhushi, Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases, *J. Pathol. Inform.* 7 (1) (2016) 29, doi:[10.4103/2153-3539.186902](https://doi.org/10.4103/2153-3539.186902).
- [16] H. Irshad, A. Veillard, L. Roux, D. Racocceanu, Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential, *IEEE Rev. Biomed. Eng.* 7 (2014) 97–114, doi:[10.1109/rbme.2013.2295804](https://doi.org/10.1109/rbme.2013.2295804).
- [17] P. Naylor, M. Lae, F. Reyal, T. Walter, Nuclei segmentation in histopathology images using deep neural networks, in: *Proceedings of the IEEE 14th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2017, doi:[10.1109/isbi.2017.7950669](https://doi.org/10.1109/isbi.2017.7950669).
- [18] F. Meyer, S. Beucher, Morphological segmentation, *J. Vis. Commun. Image Rep.* 1 (1) (1990) 21–46, doi:[10.1016/1047-3203\(90\)90014-M](https://doi.org/10.1016/1047-3203(90)90014-M).
- [19] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, 60, Association for Computing Machinery (ACM), 2017, pp. 84–90, doi:[10.1145/3065386](https://doi.org/10.1145/3065386).
- [20] P. Naylor, M. Lae, F. Reyal, T. Walter, Segmentation of nuclei in histopathology images by deep regression of the distance map, *IEEE Trans. Med. Imaging* 38 (2) (2019) 448–459, doi:[10.1109/tmi.2018.2865709](https://doi.org/10.1109/tmi.2018.2865709).
- [21] F. Xing, L. Yang, Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review, *IEEE Rev. Biomed. Eng.* 9 (2016) 234–263, doi:[10.1109/rbme.2016.2515127](https://doi.org/10.1109/rbme.2016.2515127).
- [22] M.N. Gurcan, T. Pan, H. Shimada, J. Saltz, Image analysis for neuroblastoma classification: Segmentation of cell nuclei, in: *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2006, doi:[10.1109/embc.2006.4398536](https://doi.org/10.1109/embc.2006.4398536).
- [23] P. Rosin, T. Ellis, Image difference threshold strategies and shadow detection, in: *Proceedings of the British Machine Vision Conference*, British Machine Vision Association, 1995, doi:[10.5244/c.9.35](https://doi.org/10.5244/c.9.35).
- [24] A.A. Nawandhar, L. Yamujala, N. Kumar, Image segmentation using thresholding for cell nuclei detection of colon tissue, in: *Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 2015, doi:[10.1109/icacci.2015.7275775](https://doi.org/10.1109/icacci.2015.7275775).
- [25] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Syst. Man Cybern.* 9 (1) (1979) 62–66, doi:[10.1109/tsmc.1979.4310076](https://doi.org/10.1109/tsmc.1979.4310076).
- [26] D.-C. Huang, K.-D. Hung, Y.-K. Chan, A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images, *J. Syst. Softw.* 85 (9) (2012) 2104–2118, doi:[10.1016/j.jss.2012.04.012](https://doi.org/10.1016/j.jss.2012.04.012).
- [27] X. Chen, X. Zhou, S. Wong, Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy, *IEEE Trans. Biomed. Eng.* 53 (4) (2006) 762–766, doi:[10.1109/tbme.2006.870201](https://doi.org/10.1109/tbme.2006.870201).
- [28] X. Yang, H. Li, X. Zhou, Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and kalman filter in time-lapse microscopy, *IEEE Trans. Circuits Syst. I Reg. Pap.* 53 (11) (2006) 2405–2414, doi:[10.1109/tcsi.2006.884469](https://doi.org/10.1109/tcsi.2006.884469).
- [29] H. Fatakdawala, A. Basavanhally, J. Xu, G. Bhanot, S. Ganesan, M. Feldman, J. Tomaszewski, A. Madabhushi, Expectation maximization driven geodesic active contour with overlap resolution (EMaGACOR): Application to lymphocyte segmentation on breast cancer histopathology, in: *2009 Ninth IEEE International Conference on Bioinformatics and Bioengineering*, 57 (7), IEEE, 2010, pp. 1676–1689, doi:[10.1109/bibe.2009.75](https://doi.org/10.1109/bibe.2009.75).
- [30] P. Yan, X. Zhou, M. Shah, S. Wong, Automatic segmentation of high-throughput RNAi fluorescent cellular images, *IEEE Trans. Inf. Technol. Biomed.* 12 (1) (2008) 109–117, doi:[10.1109/titb.2007.898006](https://doi.org/10.1109/titb.2007.898006).
- [31] M.E. Plissiti, C. Nikou, Overlapping cell nuclei segmentation using a spatially adaptive active physical model, *IEEE Trans. Image Process.* 21 (11) (2012) 4568–4580, doi:[10.1109/tip.2012.2206041](https://doi.org/10.1109/tip.2012.2206041).
- [32] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, A. Sethi, A dataset and a technique for generalized nuclear segmentation for computational pathology, *IEEE Trans. Med. Imaging* 36 (7) (2017) 1550–1560, doi:[10.1109/tmi.2017.2677499](https://doi.org/10.1109/tmi.2017.2677499).
- [33] H. Chen, X. Qi, L. Yu, Q. Dou, J. Qin, P.-A. Heng, DCAN: deep contour-aware networks for object instance segmentation from histology images, *Med. Image Anal.* 36 (2017) 135–146, doi:[10.1016/j.media.2016.11.004](https://doi.org/10.1016/j.media.2016.11.004).
- [34] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014).
- [35] N. Japkowicz, S. Stephen, The class imbalance problem: a systematic study, *Intell. Data Anal.* 6 (5) (2002) 429–449, doi:[10.3233/IDA-2002-6504](https://doi.org/10.3233/IDA-2002-6504).
- [36] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., Tensorflow: a system for large-scale machine learning, in: *OSDI*, 16, 2016, pp. 265–283.
- [37] H. Dong, A. Supratak, L. Mai, F. Liu, A. Oehmichen, S. Yu, Y. Guo, Tensorlayer: a versatile library for efficient deep learning development, *ACM Multimedia* (2017).
- [38] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2015, doi:[10.1109/cvpr.2015.7298965](https://doi.org/10.1109/cvpr.2015.7298965).
- [39] S. Xie, Z. Tu, Holistically-nested edge detection, in: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, IEEE, 2015, doi:[10.1109/iccv.2015.164](https://doi.org/10.1109/iccv.2015.164).
- [40] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: *Lecture Notes in Computer Science*, Springer International Publishing, 2015, pp. 234–241, doi:[10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [41] P.O. Pinheiro, T.-Y. Lin, R. Collobert, P. Dollár, Learning to refine object segments, in: *Proceedings of the Computer Vision-ECCV*, Springer International Publishing, 2016, pp. 75–91, doi:[10.1007/978-3-319-46448-0\\_5](https://doi.org/10.1007/978-3-319-46448-0_5).
- [42] S. Zafari, T. Eerola, J. Sampo, H. Kalviainen, H. Haario, Segmentation of overlapping elliptical objects in silhouette images, *IEEE Trans. Image Process.* 24 (12) (2015) 5942–5952, doi:[10.1109/tip.2015.2492828](https://doi.org/10.1109/tip.2015.2492828).
- [43] S. Zafari, T. Eerola, J. Sampo, H. Kalviainen, H. Haario, Segmentation of partially overlapping nanoparticles using concave points, in: *Advances in Visual Computing*, Springer International Publishing, 2015, pp. 187–197, doi:[10.1007/978-3-319-27857-5\\_17](https://doi.org/10.1007/978-3-319-27857-5_17).
- [44] K. He, G. Gkioxari, P. Dollar, R. Girshick, Mask r-CNN, in: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, IEEE, 2017, doi:[10.1109/iccv.2017.322](https://doi.org/10.1109/iccv.2017.322).
- [45] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, Focal loss for dense object detection, *IEEE Trans. Pattern Anal. Mach. Intell.* (2018) 1, doi:[10.1109/tpami.2018.2858826](https://doi.org/10.1109/tpami.2018.2858826).



**Lipeng Xie:** received the B.S. degree in communication engineering from Nanchang University, in 2013, and the M.S. degree from University of Electronic Science and Technology of China, Chengdu, China, in 2016. He is currently with the School of Communication and Information Engineering, University of Electronic Science and Technology of China as a Ph.D. student. His research interests include machine learning, image processing, medical image segmentation and object detection.



**Jin Qi:** received the M.Sc. degree in mathematics from Sichuan Normal University, Chengdu, China, in 2002, and the Ph.D. degree in computer science from the Institute of Automation, Chinese Academy of Science, Beijing, China, in 2005. From 2005 to 2007, he was with the Department of Electrical Engineering, University of Electronic Science and Technology of China, as an Assistant Professor and from 2007 to 2009, as an Associate Professor. In 2010, he was a Visiting Scholar with the University of Chicago. From 2011 to Oct. 2013, he was a Postdoc with the Northwestern University. From Nov. 2013 to Oct. 2014, he was with the Medical College of Georgia. Since Nov. 2014, he has been with Childrens National Medical Center as a Postdoc. His research interests include fingerprint recognition, biometrics, image processing, computer vision, pattern recognition, and machine learning.



**Lili Pan:** received her B.Eng. degree in Electronic Engineering, as well as his M.Eng. and Ph.D. degrees in Information Engineering from School of Electronic Engineering at University of Electronic Science and Technology of China (UESTC), Chengdu, China in 2004, 2007, and 2012, respectively. From 2009 to 2011, she was a visiting student in the Robotics Institute at Carnegie Mellon University, Pittsburgh, USA. She joined Department of Information Engineering at UESTC as lecturer in 2012. She is now an associate professor of UESTC. Her research interests include computer vision and machine learning.



**Samad Wali:** received the B.Sc. degree in mathematics from Forman Christian College Lahore, Pakistan, the M.Sc. degree in applied mathematics from The Islamia University of Bahawalpur, Pakistan, and the Ph.D. degree in computational mathematics from Nankai University, Tianjin, China, in June 2018. His main areas of research interest are image processing, variational methods, and numerical solution to partial differential equations. He is currently a Research Fellow in Image segmentation and medical image analysis at the School of Communication and Information Engineering, University of Electronic Science and Technology of China.