An Improved Cervical Cell Segmentation Method Based on Deep Convolutional Network

Xueli Hao, Lili Pei, Wei Li, Ying Liu, and Hao Shen

School of Information Engineering, Chang’an University, Xi’an 710064, China

Correspondence should be addressed to Wei Li; grandy@chd.edu.cn

Received 7 April 2021; Accepted 4 February 2022; Published 10 March 2022

Copyright © 2022 Xueli Hao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cervical cytology smear test is an effective method for cervical cancer early screening, and segmentation accuracy is essential for computer-aided diagnosis. In this study, an improved cervical nucleus and cytoplasm segmentation method based on a deep convolutional network was proposed. This method consisted of a cellular region proposal and pixel-level segmentation network (CRP-PSN). Data were obtained from the 2014 International Symposium on Biomedical Imaging cervical cell segmentation competition open dataset. In CRP networks, online hard example mining and soft-nonmaximum suppression algorithms were incorporated to solve problems, including background noise, impurities, and other interferences in smear images. For PSN networks, a generative adversarial network-generated adversarial network algorithm and least-squares loss function were used to generate cell region segmentation, thereby improving the cytoplasm segmentation results. Finally, the improved CRP-PSN model was analyzed and compared with other typical cervical cell segmentation methods. The experimental results showed that the proposed model could effectively improve the segmentation accuracy of cytoplasm and nuclei in cervical cytology smear images by 92% and 98.6%, respectively. These findings provided strong support for the application of this method for automated interpretation of cervical cytology smear images and improved diagnostic reliability.

1. Introduction

Cervical cytopathology tests can detect early cervical cancer and improve cure rates. Computer-aided diagnosis can overcome several problems with manual film reading. Cervical cell segmentation is the basis for subsequent quantitative and qualitative cell analyses and is a major focus of computer-assisted cervical cytopathology screening.

Artificial intelligence has recently been used in medical image analysis and segmentation. Li et al. [1] used a spatial K-means clustering algorithm to extract the initial profiles of the nucleus and cytoplasm. A stack-based differential compensation method was proposed to improve errors in segmenting cytoplasmic boundaries. Moreover, Phoulady et al. [2] found the nucleus using adaptive thresholds and then segmented cell clusters using a Gaussian mixture model with pixel-level expectation-maximization to achieve boundary approximation and boundary refinement. Ramalho et al. [3] used a clustering algorithm with a Voronoi diagram algorithm to segment cell gaps, combined with the Canny edge detection algorithm to obtain more accurate cell boundaries. In 2016, Zhao et al. [4] proposed a segmentation framework based on iterative clustering and a Markov random field algorithm. Song et al. [5] then designed a multiscale convolutional neural network that uses a priori shape template matching to set the initial profile for each cervical cell, and shape fitting was carried out by morphological methods. In 2017, Zhang et al. [6] used a fully convolutional network to obtain the results of crude cervical cell nucleus localization; changes in the grayscale value of the nucleus edge were applied to construct a loss function and perform a search of the nucleus boundary, enabling segmentation. Additionally, Pan et al. [7] improved segmentation performance by removing interfering information in the background and highlighting nuclei in pathology images through sparse reconstruction methods. A simple linear iterative clustering technique was developed by Tareef et al. [8] to divide the cell image into multiple block regions,
enabling optimization of the segmentation using Voronoi graphs and a level set based on a learning shape prior. Subsequently, Tareef et al. [9] proposed a multichannel fast watershed method to partition the nucleus and cytoplasm. Yu et al. [10] divided cervical cell smear images into several levels to achieve segmentation of most cervical cells.

In 2019, Zhao et al. [11] adopted a window segmentation approach to implement a watershed segmentation algorithm for the whole image segmentation. Li et al. [12] then used a convolutional network semantic segmentation model to implement segmentation of free nuclei and cytoplasm in images of cervical cells. In a separate study, the global threshold method with rough segmentation was improved using gradient edge information [13], yielding initial contours of nuclei; this approach improved segmentation accuracy in cervical cells by placing specific initial contours combined with the gradient vector flow (GVF) snake model. In 2019, Araujo et al. [14] introduced computational tools for cytological analysis that incorporated cell segmentation deep learning techniques, and Li [15] proposed a generative adversarial network (GAN), Cell-GAN, for cell segmentation, enabling isolation of cervical cells from the background area and from other cells outside the overlapping area to obtain individual cervical cells and information on the edge of the cells, facilitating segmentation. Notably, Han [16] combined the OTSU algorithm and mathematical morphological manipulation to detect cell clusters in images; a more accurate segmentation of the nucleus region was achieved using a Chan Vese model. In 2020, Hu et al. [17] proposed an effective approach for computed tomography lung segmentation using mask region-based convolutional neural networks.

Based on these different segmentation algorithms, segmentation of cervical cells in liquid-based thin-layer smear images can be achieved using traditional algorithm models based on thresholds, field strength, and convex and concave point curvature connections and supervised learning models from convolutional neural networks. Traditional threshold-based segmentation methods are applicable to single or nonoverlapping cytoplasm or nuclei, and the cell must have good contrast with the background. The accuracy of segmentation is affected by various factors, such as image brightness and stain distribution. In addition, there are many overlapping cells in cell smears during cell detection; therefore, the accuracy and practicality of segmentation by these methods need to be improved. Although cell segmentation methods based on the watershed, GVF snake model, and field strength can identify cell boundaries, these types of algorithms are extremely susceptible to interference from abnormal pixel values in the images. At the same time, the watershed algorithm is sensitive to the initial marker information for the target region. This approach is also prone to problems such as oversegmentation, whereas the GVF snake model often requires manually setting the starting point and is easily disrupted by impurities or false contours of overlapping areas of cells owing to factors such as overlapping and opaque cell areas faced by cervical cell image segmentation. Thus, improvements in these methods are necessary.

Accordingly, in this study, we developed an improved method for segmentation of cervical nuclei and cytoplasm in cervical cell smear images based on a deep convolutional network, i.e., a cellular region proposal and pixel-level segmentation network (CRP-PSN). This method improved diagnostic reliability and provided strong support for automated cervical cytology smear images diagnosis.

2. Material and Methods

In this study, the Tensorflow deep learning framework was used in Ubuntu16.04 for model design, development, and training. The computer was equipped with an Intel i7 CPU, 32 GB of RAM, and two GTX1080ti GPUs with 11 G video memory.

A public dataset from the 2014 International Symposium on Biomedical Imaging (ISBI) was used in this study. This dataset contained 961 cervical cell smear extended depths of field images with nuclear and cytoplasmic labels. The image composition and dataset partitioning are shown in Table 1. Image samples are shown in Figure 1.

Since the training dataset was small and contained images of different sizes, 1024 × 1024 cell images were cropped using a 512 × 512 pixel rectangle and a sliding window with a step size of 256 pixels to split the eight 1024 × 1024 images into 72 images measuring 512 × 512 pixels. In addition, we used 90°, 180°, and 270° image rotation; horizontal and vertical flipping; adding noise (Gaussian noise and salt pepper noise); image brightness; contrast adjustment; and Gamma transform to expand 207 training images six-fold to 1242 to provide data support for model training.

The CRP-PSN was proposed to make full use of the structural characteristics and spatial information of cells and to achieve accurate segmentation of cervical cancer cells. To avoid interference between the cell area detection network and the pixel-level segmentation network, the model training process was divided into two stages (CRP-PSN), as shown in Figure 2.

The CRP network could be used to localize cervical cells and allow the region to be segmented while the PSN performed pixel-level segmentation of cervical cells in the target region.

The CRP network will be less accurate if there is a large overlap of cells. The goal is to optimize and improve the CRP-PSN model. For the CRP network, online hard example mining (OHEM) and soft-nonmaximum suppression (NMS) algorithms were added to improve the ability of the approach to detect cell regions and to overlap cells. Because of the GAN algorithm and least-squares loss function for PSNs, this approach enabled cell area segmentation, thereby improving the segmentation of the cytoplasm.

2.1. Improvement of the CRP Network Based on OHEM and Soft-NMS

2.1.1. Improvement Based on OHEM. OHEM is trained in loss calculation and learning for difficult samples with misclassification during training to strengthen the classification network for difficult sample mining. The improved CRP network based on the OHEM structure is shown in Figure 3.
As shown in Figure 3, there were two identical network structures \( (a \text{ and } b) \), \( a \) propagated forward for all regions of interest (ROIs) and its weight parameters were shared in the weight parameters of \( b \). The \( b \) network performed both forward and backward propagation of ROIs filtered by \( a \), and the parameters were continuously optimized through iterative training.

The OHEM structure was propagated forward for all ROIs, ranking the ROIs based on categorical and regression loss values. The top 256 ROIs were selected for backpropagation. The loss values of each ROI were calculated and sorted, and the NMS algorithm was used to remove the overlapping ROIs from the sample and then select the 256 ROIs for backpropagation.

One iteration training process of OHEM was as follows: compute the feature map, perform forward calculations on all ROIs, and compute the loss of each ROI. The NMS and sorting unit were used to select ROIs with higher loss values. These ROIs were entered into a readable and writable ROI network to perform forward calculation and backpropagation to update the network. The parameters of the readable and writable ROI network were assigned to the readable-only network to complete an iteration.

2.1.2. Improvement Based on Soft-NMS. The NMS algorithm was applied in the CRP network to eliminate redundant cross-repeat windows and preserve the optimal cell detection bounding box. NMS forces the confidence level of all ROIs with overlapping proportions greater than the unique threshold to zero, as shown in the following equation:

![Figure 1: Data sample: (a) 1024 × 1024 pixel and (b) 512 × 512 pixel data sample.](image1)

![Figure 2: CRP-PSN model structure.](image2)

![Figure 3: Flow chart of the OHEM structure.](image3)

**Table 1: Image composition and data set partitioning.**

<table>
<thead>
<tr>
<th>Image type</th>
<th>Image number</th>
<th>Image pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic image</td>
<td>945</td>
<td>512 × 512</td>
</tr>
<tr>
<td>Real image</td>
<td>16</td>
<td>1024 × 1024</td>
</tr>
<tr>
<td>Training set</td>
<td>135</td>
<td>512 × 512</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1024 × 1024</td>
</tr>
<tr>
<td>Validation set</td>
<td>810</td>
<td>512 × 512</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1024 × 1024</td>
</tr>
</tbody>
</table>
\[ s_i = \begin{cases} s_i, & \text{iou}(M, b_i) < N_r, \\ 0, & \text{iou}(M, b_i) \geq N_r. \end{cases} \] (1)

If the proportion of overlap between the two cell regions is large, it will cause one of the cells to fail detection, and the average detection rate of the CRP network will be reduced. As shown in Figure 4, the detection network should output two boxes; however, because the intersection over union (IOU) ratio between the a and b detection boxes was significantly greater than 50%, the NMS excluded the detection frame with low confidence, resulting in cells in the region being missed.

To solve this problem, we used the soft-NMS algorithm to set confidence levels for ROIs that had overlapping regions for improvement of the threshold-only direct rejection drawback of the NMS algorithm. As shown in equation (2), if two ROIs had a large proportion of overlap, it was appropriate to reduce the confidence level of one of the ROIs rather than directly censoring to obtain a balance between the threshold and confidence in NMS.

\[ s_i = \begin{cases} s_i, & \text{iou}(M, b_i) < N_r, \\ s_i(1 - \text{iou}(M, b_i)), & \text{iou}(M, b_i) \geq N_r. \end{cases} \] (2)

Depending on the proportion of overlapping detection regions, the confidence of the detection frame was attenuated by (1–IOU_Ratio), assuming that the overlap area between the detection frame and the b detection frame was 60% of the detection frame; the confidence level of the detection frame was multiplied by an attenuation factor of 0.4 (0.85 × [1 – 0.6] = 0.34), establishing a reasonable confidence threshold to retain this detection frame (Figure 5).

The use of an improved CRP network based on the soft-NMS network could reduce the number of cells missed due to overlapping to some extent, and the method was simple. However, application of the soft-NMS network will require the identification of a more appropriate confidence threshold through dataset testing.

2.2. Improvement of the PSN Based on GAN. To improve the segmentation accuracy, a GAN-based improved segmentation model was proposed. GAN’s (generative adversarial network) main inspiration comes from zero-sum game theory. It includes two models: generative model G and discriminative model D. The G captures the data distribution, and the D estimates the probability that a sample came from the training data rather than G. The structure of the model is shown in Figure 6. The model input was an image of the ROI of cervical cells determined by the CRP network, and the output was either 0 or 1. During the propagation of the GAN segmentation model, the output of the generator was a segmentation rendering having the same size as the input image. The input of the discriminator was the segmentation-rendering image of the generator and the segmentation label corresponding to that cell. The discriminator evaluated the segmentation effect of the generator by performing feature extraction and comparative analysis of the two input images.

The model had several advantages. Primarily, the inclusion of a discriminator network structure enabled input of both the segmentation effect map generated by the generator and the real analysis result map into the convolutional neural network. The input convolutional neural network could be used to directly calculate the loss of the generator via the network judgment function instead of the loss function. The model structures of PSN and GAN-PSN are shown in Figure 7.

The PSN relied entirely on the loss function to guide the convolutional network and approximate the real image data distribution, and the training process of the PSN involved making the input and output of the network as similar as possible. Thus, the generated images only mimicked the training set and could not generate cell regions that were heavily occluded, as shown in the following equation:

\[ L_{\theta,\phi}^{\text{VAE}} = -E_{q\phi(z|x)}[\log(p(x|g_\theta(z)))] + KL(q\phi(z|x)||p(z)). \] (3)

In GAN-PSN, the loss of information from PSN was related partly to the discriminator’s true-false results and partly to the segmentation-dependent pixel loss with real labels. Instead of using the pixel-loss function alone, a zero-sum game was played between the discriminator D and generator G. On the one hand, generator G aimed at generating a false sample (loss evaluation), spoofing discriminator D into mistaking it for the real sample, as shown in the following equation:

\[ L_G^{\text{GAN}} = E[\log(D(G(z))]]. \] (4)

On the other hand, discriminator D should have the ultimate goal of distinguishing the true sample x from a false sample G(z) (loss evaluation), as shown in the following equation:

\[ L_D^{\text{GAN}} = E[\log(D(x))]+E[\log(1-D(G(z))]]. \] (5)

The structure of the discriminator network for the GAN-improved PSN is shown in Figure 8.

As another advantage of the model, PSN used a cross-entropy loss function and an improved segmentation model based on the GAN, incorporating a least-squares loss function, as shown in the following equations:

\[ L_s(D) = \min\left\{ \frac{1}{2}E_{x\sim P_{\text{data}}(x)} \left[ (D(x) - b)^2 \right] + \frac{1}{2}E_{z\sim P_z(z)} \left[ (D(G(z)) - a)^2 \right] \right\}, \] (6)

\[ L_s(G) = \min\left\{ E_{z\sim P_z(z)} \left[ (D(G(z)) - c)^2 \right] \right\}. \] (7)
where $G$ is the generator, $D$ is the discriminator, $z$ is the inter-image data, $P_{\text{data}}(x)$ is the probability distribution for the segmentation label obedience, and $P_z(z)$ is the distribution for the generator segmentation result map data.

As shown in Figure 9, compared with cross-entropy, the least-squares loss saturated at only one point. Thus, the network generation target value constantly approached the true value, showing a good ability to optimize the samples on the classification zero boundaries and thereby facilitating the generator to better learn the features of the target region and improve the segmentation effect.

3. Model Training

Figure 10 shows the image labeled $512 \times 512$ pixels and the artificially labeled rectangular detection box from the training of the CRP model.
The cell area segmentation network utilized pixel-level segmentation. The output of the segmentation network was a dual-channel image of nucleus and cytoplasm. The PSN contained two binary images of the cytoplasmic region and the nuclear region. The PSN training images and labels are shown in Figure 11.

3.1. CRPN Model Training. During CRPN training, the Resnet ImageNet pre-training model was used to set the batch size to 4 with the adaptive moment estimation (Adam) optimizer. The training process traversed the training set 15 times with an initial learning rate of 0.0025, and the learning rate was reduced after the dataset was traversed for the eighth and 12th times. The training process required 4 h to train the network $5.4 \times 10^3$ times. The loss function and learning rate curves for the training process are shown in Figure 12.

From the loss function curve, we found that the improved CRP network reached a more stable state later in the training process, and the decline in the loss value flattened out; thus, the model converged to a more stable state. Compared with the pre-improved CRP network training process, the pre-loss function value decreased faster, indicating that the improved CRP network facilitated the extraction of cell region features and converged more easily. The CRP network was tested using weights obtained from model training, and when a cell smear image was entered, the CRP network output box coordinated for each individual cell and for confidence levels. Figure 13 shows the effects of cell localization. Each generated coordinate frame contained only one complete cell.

Based on the output information from the detection network, a single-cell image to be segmented could be obtained by slice copying a single-cell image from the original image (Figure 14).

Different IOU thresholds and confidence thresholds can affect the number of detection regions in the model output. To determine the optimal threshold, the F1 value is used as a reference. According to the detection results from the detection model, the IOU thresholds and confidence thresholds were used to traverse the preimprovement and postimprovement detection models. According to the test results of the threshold traversal, before the CRP network improvement, the F1 maximum was 0.9505, corresponding to a confidence threshold of 0.4, an IOU threshold of 0.5, an accuracy of 0.9597, and a recall of 0.9415. The F1 maximum value of 0.9820 for the improved CRP network detection model corresponded to a confidence threshold of 0.5, an IOU threshold of 0.6, accuracy of 0.9820, and recall of 0.9820. From the model test statistics, we found that the maximum value of the improved detection model F1 increased by 3.1%. Under these conditions, the accuracy
improved by 2.2%, the recall improved by 4%, and the confidence level and IOU threshold improved by 0.1%. Thus, the CRP network improved could effectively improve the recall of the detection model and the accuracy and confidence in the detection region of the model.

3.2. GAN-PSN Training. In the training process of the GAN-PSN, the learning and training processes for generators and discriminators were alternately iterative. A batch of cell area images was input, and the segmentation effect was first generated by PSN; then, two subsequent steps were performed. In the first step, the segmentation results and images of the segmented label were jointly input into the discriminator, which performed the least-squares loss calculation on the segmentation result. The value of the loss function was propagated backward to the discriminator for parameter optimization. At this point, the discriminator was in the training mode. In the second step, the discriminator was in inference mode, and the segmentation result was entered. The segmented label image was then entered into the discriminator after the first step of optimization. The
The discriminator performed least-squares loss calculation on the segmentation result and propagated the loss function value backward to the generator for parameter optimization. These two steps were cycled iteratively to complete the training process for the GAN-PSN, as shown in Figure 15.

Using the Adam optimizer, the batch size was set to 8. The training process also traversed the training set 15 times, with an initial learning rate of 0.0005. The learning rate was reduced when traversing the dataset for the eighth and 12th times. The training process required 5 h to train the network.
The loss function and learning rate curves for the training process are shown in Figure 16.

From the loss function curves, we found that the loss function value of the generator gradually decreased during the training process using the GAN-improved PSN. The loss function value of the discriminator also decreased, and the accuracy of the pixel segmentation for the cell region generated by the generator was also smoothened, causing the model to converge to a more stable state. The GAN-PSN was tested using the weights obtained from the model training. When an image of a cell region slice was input, the network outputs the segmentation result of that cell (Figure 17).

The optimal F1 values of the segmentation network for segmentation of the nuclear region before and after improvement were 0.9383 and 0.9404, respectively, and the corresponding thresholds were 0.4 and 0.5, respectively. The improved segmentation network did not significantly enhance the nuclear region. This was partly because the nucleus was relatively independent, and the regions barely overlapped and partly because there were obvious differences among the nucleus, cytoplasm, and background.
pixels, which were easy to classify. Therefore, the segmentation network had less capacity for improvement with regard to the effects of nucleus segmentation. For the cytoplasmic region, the optimal F1 values before and after the improvement were 0.9051 and 0.9204, respectively, and the F1 value improved by 1.5%, whereas the optimal confidence level was improved by 0.1%. This demonstrated that an improved segmentation model based on GAN could improve the pixel segmentation of cytoplasmic regions.

A sample of the cervical cell segmentation results obtained in this study is shown in Figure 18.

4. Discussion

4.1. Model Evaluation Indexes. To test the accuracy and validity of the model, the results were evaluated using the precision, recall, and F1 values of the target detection indicators. The recall rate represents the ratio of the number of cell areas detected by the detection network to all real areas to be detected, and the accuracy rate represents the proportion of detected cell areas that are real and valid areas, as shown in the following equations:

\[ P = \frac{TP}{TP + FP} \]  \hspace{1cm} (8)

\[ R = \frac{TP}{TP + FN} \]  \hspace{1cm} (9)

where \( P \) is precision, \( R \) is recall, \( TP \) represents real positive, \( TN \) represents real negative, \( FP \) is false positive, and \( FN \) is false negative.

In general, higher IOU threshold and confidence threshold settings are associated with higher accuracy and lower recall of the detection model. To find a reasonable threshold, a combined assessment of accuracy and recall was performed using F1 values, as shown in equation (10); higher F1 values represent the detection capacity of the model.

\[ F1 = \frac{2P \times R}{P + R} \]  \hspace{1cm} (10)

In addition, for comparison with the 2014 ISBI evaluation indicators, the ZSI index and DC coefficients were used for the evaluation of segmented areas, as shown in the following equations:

\[ ZSI = \frac{2TP}{2TP + FP + FN} \]  \hspace{1cm} (11)

\[ DC = \frac{2|A \cap B|}{|A| + |B|} \]  \hspace{1cm} (12)

\( A \) and \( B \) in the DC coefficients represent the segmentation area label matrix and classification prediction moments, respectively, and the ZSI and DC coefficients are essentially the same as the F1 values.
4.2. Analysis of Experimental Results. To verify the validity of the design model in this study, the segmentation statistical indicators discussed by Lu et al. [18] were used here. They made a comparison of the detection results of Masoud [19], Ushizima [20], and their own results from the same dataset (2014 ISBI). Ushizima et al. [19] achieved segmentation mainly by threshold segmentation combined with Voronoi plots, where the cellular region was surrounded by polygons and the nucleus was set as the initial seed point, with a region-based growth approach. By contrast, Masoud et al. [19] combined signed-distance-field, random forest, and multiple constraints to segment cells. Additionally, Lu et al. [18] used a joint because the two categories of cellular and noncellular regions were evaluated by the PSN classifier using the sigmoid activation function; the effect drawing of the segmentation produced by the segmentation network belongs to the category probability graph of the pixel. The probability of the cytoplasmic region is biased towards 1, and the probability of the background region is biased towards 0. The results are usually binarized using a threshold of 0.5 to distinguish exactly between the cytoplasmic and background regions. However, the actual training effect of the model may be shifted by approximately 0.5. In laboratory studies, the choice of the threshold value can affect the segmentation effect and test index to some extent. Therefore, the cytoplasmic and nuclear channels in the segmentation results were binarized and tested with different thresholds (0.2, 0.3, 0.4, 0.5, 0.6, and 0.7). Optimization of the method for segmentation was based on multilevel set functions, such as cell contour length, edge intensity, and area. These three representative traditional methods have already achieved good segmentation results in this dataset. Figure 19 shows the segmentation effects of this method and the other three methods in the same image. The method of Ushizima et al. [20] did not establish precise polygon boundaries, and both our current method and the segmentation renderings of Masoud et al. [19] and Lu et al. [18] showed different degrees of error in the segmentation process. However, the currently proposed method was superior with regard to cell boundaries, particularly for overlapping boundaries where contrast is poor.

A quantitative comparative analysis of the results of the four segmentation methods is shown in Tables 2 and 3.
Segmentation capabilities through supervised learning. 

A model with better cell region detection and complexity of the segmentation algorithm design and using convolutional feature extraction also simplified the methods. At the same time, the deep learning approach segmentation model proposed in this study over traditional methods. Through comparative analysis, the relevant metrics data showed significant advantages when using the deep learning-based cell segmentation model proposed in this study over traditional methods. At the same time, the deep learning approach using convolutional feature extraction also simplified the complexity of the segmentation algorithm design and provided a model with better cell region detection and segmentation capabilities through supervised learning.

### 5. Conclusion

The results of training and test trials on the improved CRP-PSN model showed that soft-NMS and OHEM could improve the completeness and confidence of CRP-PSN in the detection of cell areas, particularly for cervical cells with large overlapping areas and more complex cell regions. The segmentation accuracy of the GAN-improved segmentation network was also improved for cellular regions, resulting in a reduction in false segmentation of overlapping cellular regions and improvement of the segmentation of cytoplasmic and nuclear boundaries. This demonstrated that the PSN was improved with the GAN method, which was effective for pixel segmentation of cellular regions. In future studies, the size of the training dataset can be further increased to improve the robustness of the model. Furthermore, in cell area detection, the existing rectangular detection area can be replaced by an angular boundary box or ellipse detection box to further improve the location accuracy of cell area. In terms of pixel segmentation, more efficient network structures, loss functions, and optimization functions can be designed to further improve the accuracy and efficiency of the detection model and segmentation model.

### Data Availability

Public dataset from the 2014 International Symposium on Biomedical Imaging (ISBI) was used in this study. This dataset contained 961 cervical cell smear extended depth of field images with labeled nuclei and cytoplasmic labels.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

The work was supported by the National Natural Science Foundation of China (51908059 and 51978071), Fundamental Research Funds for the Central Universities, CHD (300102242006 and 300102249301).

### References


### Table 2: Comparison of nucleus segmentation results.

<table>
<thead>
<tr>
<th></th>
<th>Precision (object)</th>
<th>Recall (object)</th>
<th>Precision (pixel)</th>
<th>Recall (pixel)</th>
<th>ZSI (pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ushizima et al. [20]</td>
<td>0.959</td>
<td>0.895</td>
<td>0.968</td>
<td>0.871</td>
<td>0.914</td>
</tr>
<tr>
<td>Masoud et al. [19]</td>
<td>0.903</td>
<td>0.893</td>
<td>0.901</td>
<td>0.916</td>
<td>0.9</td>
</tr>
<tr>
<td>Lu et al. [18]</td>
<td>0.977</td>
<td>0.883</td>
<td>0.942</td>
<td>0.912</td>
<td>0.921</td>
</tr>
<tr>
<td>CRP-PSN</td>
<td>0.986</td>
<td>0.992</td>
<td>0.902</td>
<td>0.982</td>
<td>0.940</td>
</tr>
</tbody>
</table>

### Table 3: Comparison of cytoplasmic segmentation effects.

<table>
<thead>
<tr>
<th></th>
<th>DC</th>
<th>FNo</th>
<th>TPp</th>
<th>FPp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ushizima et al. [20]</td>
<td>0.872</td>
<td>0.267</td>
<td>0.841</td>
<td>0.002</td>
</tr>
<tr>
<td>Masoud et al. [19]</td>
<td>0.871</td>
<td>0.111</td>
<td>0.875</td>
<td>0.004</td>
</tr>
<tr>
<td>Lu et al. [18]</td>
<td>0.893</td>
<td>0.316</td>
<td>0.905</td>
<td>0.004</td>
</tr>
<tr>
<td>CRP-PSN</td>
<td>0.920</td>
<td>0.126</td>
<td>0.913</td>
<td>0.002</td>
</tr>
</tbody>
</table>


