IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 11, 2021, accepted August 10, 2021, date of publication August 13, 2021, date of current version August 25, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3104609

Segmentation of Cervical Cell Images Based on Generative Adversarial Networks

JINJIE HUANG[1,2], GUIHUA YANG[1,3], BIAO LI[2], YONGJUN HE[1], AND YANI LIANG[2]

[1]School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China
[2]School of Automation, Harbin University of Science and Technology, Harbin 150080, China
[3]School of Electrical and Mechanical Engineering, Daqing Normal University, Daqing 163712, China

Corresponding author: Jinjie Huang (huangjinjie163@163.com)

This work was supported in part by Heilongjiang Provincial Key Laboratory of Complex Intelligent System and Integration, and in part by the Natural Science Fund Project of Heilongjiang Province of China under Grant F201222.

ABSTRACT The segmentation of cervical cell in liquid-based smear image plays an important role in cervical cancer detection. Despite of research for many years, it is still a challenge for the complexity of cell images such as poor contrast, cell irregularity, and overlapping. To solve this problem, a novel method is proposed based on Cell-GAN - a generative adversarial network. Firstly, the Cell-GAN is trained to learn a probability distribution of cell morphology by comparing the difference between the generated single-cell images and annotated single-cell images. Thus, the Cell-GAN has the ability to judge the integrity of a cell and treat other cellular information of a cell image, except for overlapping parts, as the background. Then, a complete single-cell image is generated by the trained Cell-GAN for each cell, which is located by a guide factor. The guide factor is constructed by a part of the cell to be segmented, such as the nucleus, to help Cell-GAN locate the cell and avoid generating a multi-cell image in the presence of overlapping, which means the contours of cells still cannot be distinguished. Finally, the segmentation line is defined by the contour of the generated cell, and the input image is cropped using the cell size information. The cropped image is reused for image generation until the area of generated cell varies within a small range.The proposed method is evaluated on the segmentation of single-cell images and overlapping cell images and obtained significant values of 94.3% DC, 7.9% FNR_o for single-cell images and 89.9% DC, 6.4% FNR_o for overlapping cell images respectively. The experimental results indicate that the proposed method can adaptively approach the boundary lines of cells to handle with different cases of overlapping in cervical cell images through what learned by the Cell-GAN. The proposed method outperforms most current methods in both segmentation accuracy and robustness.

INDEX TERMS Deep learning, Generative Adversarial Networks, cervical cell image segmentation, overlapping cells.

I. INTRODUCTION

Cervical cancer is the fourth most common cancer in women worldwide [1].According to Global Cancer Observatory (GCO) in 2018, there is an estimated 570,000 new cases and 311,000 deaths due to cervical cancer [2]. According to GCO latest data, there is an estimated 604,127 new cervical cancer cases in 2020. Early diagnosis and early treatment are considered as the only effective way to reduce the risk of cervical cancer [3]. At present, the Thinprep cytologic test (TCT), combined with the human papillomavirus (HPV) detection, significantly improves the screening sensitivity and specificity for cervical cancers, which is currently the most effective screening method for the disease [4].This method requires an experienced pathologist to make a diagnosis after observing the diseased nuclei under a microscope, which ties up a great deal of manpower and material resources. It is such a heavy workload for a pathologist to search several abnormal cells from more than 300 thousands of cells that it is inevitable for the pathologist to make error in diagnosis. Hence, with the development of image processing, some automatic screening techniques are used for cancer cells detection [5]. These techniques usually segment cell images at first, and then classify the segmented cells into different types, and even

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Khurram Khan.

identify those cancerous cells [6]. An effective cell segmentation method can reduce the complexity of the classifier to be trained and help to improve the diagnosis accuracy significantly.

So far, cell image segmentation methods can be grouped into two categories: traditional cell segmentation methods and deep-learning-based cell segmentation methods [7]. Traditional cell segmentation methods include watershed algorithm [8], region growing segmentation [9], level set [10], and joint optimization methods of multiple level set functions [11]. N. Mat-Isa *et al.* [12] proposed a method named the seed region growing features extraction (SRGFE) to segment single-cell images. T. R. Jones *et al.* [13] improved region growing method based on computing an approximation of the Voronoi regions of each seed on a manifold to segment overlapping cell images. However, this method could not deal with poor contrast cell images and could not extract cells from the background, because the separation of foreground and background was not considered there. A. Kale *et al.* [14] combined hierarchical trees built by the information of spectrum, shape, and gradient of cells with multiple watershed algorithms to segment poor contrast and overlapping cells. Nevertheless, over-segmentation of cells appeared serious when the overlapping degree was high in this method. Z. Lu *et al.* [11] modified multiple level set functions by optimizing the energy function based on contour length, edge strength, prior knowledge of shape, and area of overlapping, to segment overlapping cell images. But the segmentation performance is still poor for those cell images with complex background and poor contrast. Phoulady *et al.* [15] detected candidate nuclei based on iterative thresholding. Thresholding based segmentation often fail for images with cell overlap, where overlapping cytoplasm parts appear with similar mean intensity as nucleus.

Cell segmentation methods based on deep learning include two frameworks: convolutional neural networks (CNN) [16] and fully convolutional networks (FCN) [17]. Song Y *et al.* [18] used CNNs to segment the cytoplasm of cervical cells for the first time. And then, they proposed to use multi-scale CNNs to segment cervical cell images. However, this method can only get the outer contour of overlapping cells. In order to segment overlapping cell images, it is necessary to combine CNNs with traditional segmentation methods, such as level set [19]. O.Ronneberger *et al.* [20] proposed U-Net for biomedical image segmentation and got the best ranking on the EM segmentation challenge. Although the U-Net has good segmentation performance on single-cell images, but it does not work well on overlapping cells [21]. There assume that the segmentation between background and cells can be regarded as a binary classification problem of pixels, but it is difficult for U-Net to classify the overlapping areas of cells into two or more cells at the same time [22]. As a result, the researchers turn attention to generative adversarial networks (GAN), to segment cell images by image generation rather than pixels classification.

Generative Adversarial Networks(GAN) [23] have been proposed by Goodfellow in 2014. In a simple GAN, the generator takes a random noise vector as an input and outputs an image [24]. The original purpose of GANs is to generate new images based on existing images [25]. Then they have been widely used in fields of image inpainting [26], [27], super-resolution image reconstruction [28], and image synthesis [29], [30].

In this paper, a novel segmentation method based on deep learning technique and generative model is proposed to segment overlapping cervical cell images. The highlight of the proposed method is to segment both single-cell images and overlapping cell images effectively. It can better capture the topological relationship between pixels, whether local disturbance changes (such as overlapping) or global smooth changes (such as poor contrast), by using convolution structure [31]. It has the following advantages:

A. UTILIZE MULTI-INFORMATION OF THE LOCAL, STRUCTURAL, AND PROBABILITY DISTRIBUTIONAL

As a generation model, the segmentation mechanism of Cell-GAN is different from CNN and FCN. Compared to the CNN and FCN, both of which just use local image information for pixel classification, Cell-GAN utilizes the structure information of the entire input image and the probability distribution of cell morphology to segment cells. Hence Cell-GAN gets better segmentation performance, especially on highly overlapping and poor contrast.

B. COMPLETE THE COVERED CONTOUR BY PROBABILITY DISTRIBUTION

Since the contours of overlapping cells are difficult to identify, those non-overlapping cells with identifiable contours are used for data annotation. The training set contains only a few highly overlapping cells, it can mainly use non-overlapping cells to learn the probability distribution of contour shapes of cells. Naturally, a reasonable part of contour can be provided for the part of a covered overlapping cell.

C. POSSESS A STRONG ROBUSTNESS

For traditional methods, cervical cell segmentation is seriously dependent on its nuclear segmentation following that each nucleus corresponds to a cell. So, if a nuclear is wrongly segmented, the cervical cell segmentation will be misled in traditional methods. Hence, nuclear segmentation needs to be as accurate as possible. But for the proposed Cell-GAN method, cervical cell segmentation is mainly dependent on the guide factor, while the guide factor can be served by not only its nucleus but also impurities such as inflammatory cells in the image. That is to say, if an impurity is mistakenly referred to as a nucleus, which is usually called pseudo nucleus, but as a guide factor, it doesn't affect the cervical cell segmentation result of Cell-GAN. Therefore, the proposed Cell-GAN method has a strong robustness.

II. RELATED WORK

Segmentation is one of the most effective and extensive research fields in medical image analysis. The main goal of segmentation is to partition the image into multiple regions to quickly analyze the cell [32].Numerous works have been performed on automatic segmentation of cytoplasm and nucleus of a cervical cell. In [18], a super-pixel and CNN based segmentation methods are proposed for cervical cell segmentation. The author performs the cytoplasm segmentation first since the background, and cytoplasm contrast is not evident. Then CNN is applied here to detect the region of interest. They achieved an F1 score above 0.89. In [33], the author has developed a deep learning method via a multiscale convolutional neural network (MSCN) for feature extraction and graph partitioning for nucleus segmentation. The MSCN and graph partitioning algorithms can separate cytoplasm, nuclei, and background. However, to solve the overlapping nucleus in the cell image, a new robust nuclei clump algorithm is introduced. In [19], the author proposed a unique deep CNN-based framework that can accurately segment the cervical cells form overlapping clumps. They achieved the dice coefficient (DC) value 0.91 for cytoplasm and 0.93 for nucleus in the ISBI challenge dataset and 0.90 and 0.92 for cytoplasm and nucleus respectively in the private dataset. The author in [34] presents a robust variational segmentation framework based on pixel-wise CNN and a learned shape prior, that can successfully segment nuclei and cytoplasm form densely overlapping mass. The proposed method first classifies the cellular components into the background, nuclei, and cytoplasm using CNN. Later, individual cytoplasm segmentation work is performed with Voronoi segmentation and dynamic shape prior based level set evaluation. They achieved ZSI value 0.90, TP value 0.95 for cytoplasm segmentation. The author in [35] presents an approach to segment the whole cervical cell image by using Mask R-CNN and transfer learning. ResNet10 is the backbone of Mask R-CNN. In the beginning, the cell areas are partitioned by using MASK R-CNN. Later, the cell areas are classified into nucleus and cytoplasm. The proposed algorithm produced a precision, recall, and ZSI value of 0.92, 0.91, and 0.90, respectively. In [36], A novel Instant Relation Network (IR-Net) is advised to segment the overlapping cervical cell. ResNet-50 is the backbone of the IR-Net. They achieved Average Jaccard Index (AJI) and F1 value of 0.7185 and 0.7497 for cytoplasm and 0.5496 and 0.7554 for nucleus segmentation.

It is observed that most of the research work for the segmentation of overlapping cells has been conducted using the CNN or FCN, minimal effort has been made so far with Generative Adversarial Networks(GAN). In [37], the authors replace the traditional discriminator with a fully convolutional multiclass classifier. The classifier assigns to each input image pixel one label that corresponds to a semantic class or to fake/real mark. In this way, they use unlabeled non-real images created through Generative Adversarial Networks during the training process. In [38], the discriminator is adapted to distinguish between manually segmented cell microscopy images and generated images from CNNs. The generated (estimated) segmentation images are similar to manually annotated images and therefore are more accurate than those obtained from a simple CNN segmentation model. In addition, such methods have been used as domain adaptation techniques [39] to transform magnetic resonance images (MRI) into computed tomography (CT) images [40] or Differential Interference Contrast to Phase Contrast microscopy images [41]. Inspired by these applications, the researchers assume that segmentation of cell images can be regarded as a problem of filtering of invalid image information, especially under the constraint of single-cell generation, the rest of image information can be considered invalid except for a special cell. Therefore, the main task is to design a model that can learn the probability distribution of cell morphology, which is what GAN is good at.

III. METHOD

Segmentation process for cervical cell image in this work can be divided into two steps, as shown in Figure 1. The first step is to generate a segmented cell image with an approximate contour to the cell in the input image. For this purpose, the Cell-GAN is designed based on generative adversarial networks [23]. The Cell-GAN consists of a generator and a discriminator. The generator is used to generate segmented cell images. The discriminator is applied to help generator to learn the probability distribution of cell morphology by evaluating the difference between the segmented cell images and the annotated cell images. Furthermore, L_2 loss function is introduced to capture the overall structure of cell images and to adjust the generating scale of Cell-GAN for uncertain cell area, retained or discarded. In this step, to ensure that the generated results can be used to segment cells, the researchers limit that only one cell can be generated in each generated image, which is achieved with the help of guide factors. The constraint makes Cell-GAN focus on generating a complete cell whose contour is close to the cell to be segmented. The other cellular information of input cell image, except for overlapping parts, is treated as background. The guide factor is constructed by a part the cell to be segmented, such as the nucleus, and used to locate the cell to be segmented. Without the help of guide factors, overlapping and multi-cell images would make Cell-GAN so puzzled that could not figure out which cell should be processed and a multi-cell image would be generated, whose contours are still hard to be distinguished. More details of guide factor are given in section 3.3.1.

The second step is to extract real segmented cell image. Because there is noise around the generated cell sometimes, denoising is necessary before extraction. Under the constraint of single-cell generation, it's found that the generated cell has the larger contour than noise in the generated image. Therefore, the researchers combine the maximum contour search with open operation and close operation to complete

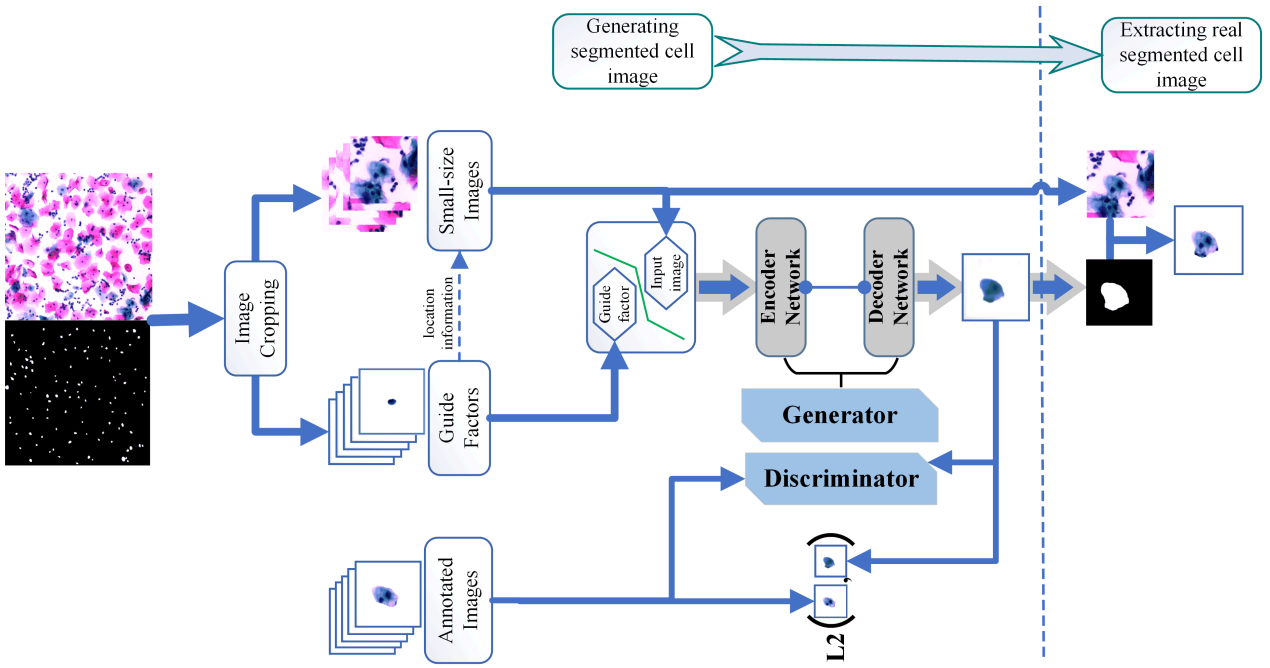


FIGURE 1. Panorama of the segmentation process for cervical cell image.

the transformation of the generated cell image to the segmented cell image. At this point, segmentation is finished.

A. CELL-GAN

GAN models can be considered as a two-player game between a generator, which learns how to generate samples resembling real data, and a discriminator, which learns how to discriminate between real and generated data. Both the generator and the discriminator cost functions are minimized simultaneously. The iterative minimization of cost functions eventually leads to a Nash equilibrium where neither can further unilaterally minimize its cost function. In the end, the GAN discriminator provides an abstract unsupervised representation of the input images. Cell-GAN is derived from DCGAN [24]. Traditionally the target of GANs [23] is to convert the random noise z into an image x by generator G, mapped by G: $z \rightarrow x$. Some improved GANs, for example [26], try to repair incomplete image $\tilde{x}$ conditioned on its surroundings, defined by G: $\tilde{x} \rightarrow x$, where x is the complete image.

In this work, Cell-GAN is contrary to image inpainting intuitively and attempts to generate single-cell image $\hat{x}$ based on cell image x, which can be described as: G: $x \rightarrow \hat{x}$. To fulfill this task, the researchers redesign the generator based on auto-encoder [42] and inception model [43], though the discriminator has the similar architecture as the original DCGAN.

1) GENERATOR

The generator consists of an encoder network and a decoder network. A schematic view of the generator is depicted in Figure 2. The encoder network receives two inputs at the same time, one is the cell image with background, and the other is the guide factor that is constructed by the nucleus to help Cell-GAN to locate the cell to be segmented. Considering the balance between the execution speed and the segmentation accuracy, the sizes of inputs are set as 200 × 200 pixels, RGB color channels is used for mean subtraction [44]. Furthermore, to establish the connection between the two inputs, two convolutional layers are used to deal with the guide factor specifically, and then merge the computed feature maps into encoder network by additive operation [45].

The main structure of the encoder network is a four-layer down-sample network, and per-layer down-sample network adopts inception architecture [43]. To promote the generalization ability of the encoder network, the sizes of groups of filters are 1 × 1, 3 × 3, and 5 × 5. According to the guidelines for stable DCGANs [24], the max-pooling layer and add Leaky-ReLU activation for each group of filters in the inception architecture are ignored. In addition, use 3 × 3 convolutions with stride 2 to replace max-pooling layer used for dimension reduction.

In general, the traditional auto-encoder uses the full connecting to realize the connection between the encoder output and the decoder input, both of which have the same sizes of feature maps. However, in proposed architecture, given the output of encoder network is $4 \times 4 \times 512 = 8192$, the number of parameters is explosive under full connection approach (over 100M). Considering that the number of 5 × 5 convolutions in encoder network can be expensive with a deeper network, the encoder network is not deepened. D. Pathak *et al.* proposed channel-wise fully connected layer so that the number of parameters can be reduced by 500 times effectively, compared with full connecting [26]. But in proposed work,

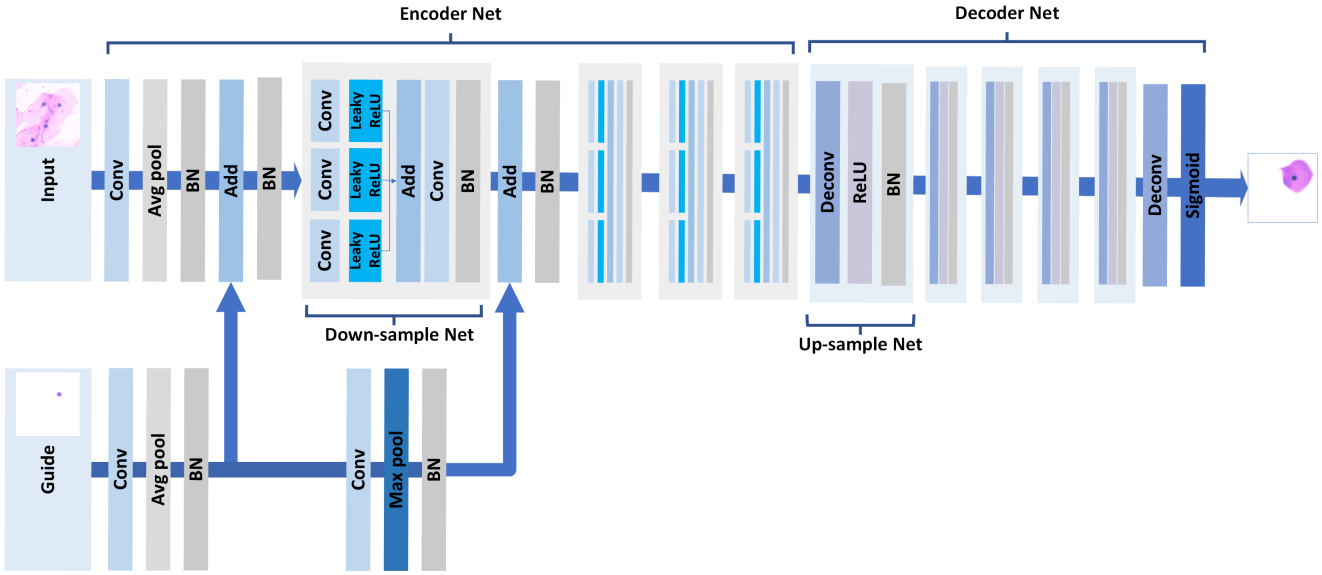


FIGURE 2. Generator network.

TABLE 1. Parameter configuration of generator.

type		patch size/stride	output size	dropout	params	ops
G – Conv 1		5 × 5 / 2	50 × 50 × 32		2.4K	0.01M
G – Conv 2		3 × 3/2	25 × 25 × 64		18.4K	0.07M
E – Conv 0		5 × 5 / 2	50 × 50 × 32		2.4K	0.01M
E – Conv 1	1 - 1	1 × 1 / 1	50 × 50 × 32		54.5K	0.22M
	1 - 2	3 × 3 / 1	50 × 50 × 32			
	1 - 3	5 × 5 / 1	50 × 50 × 32			
	conv	3 × 3 / 2	25 × 25 × 64	0.5		
E – Conv 2			13 × 13 × 128	0.5	217K	0.87M
E – Conv 3			7 × 7 × 256	0.5	868K	3.5M
E – Conv 4			4 × 4 × 512	0.5	3473K	13.9M
D – Conv 1		3 × 3 / 2	7 × 7 × 256	0.5	1179K	4.7M
D – Conv 2		3 × 3 / 2	13 × 13 × 128	0.5	295K	1.18M
D – Conv 3		3 × 3 / 2	25 × 25 × 64	0.5	73.7K	0.3M
D – Conv 4		3 × 3 / 2	50 × 50 × 32	0.5	18.4K	0.74M
D – Conv 5		3 × 3 / 2	100 × 100 × 16		4.6K	0.018M
D – Conv 6		3 × 3 / 2	200 × 200 × 3		0.4K	0.002M

G stands for guide, E stands for Encoder Net, D stands for Decoder Net

the integrity of the contour is more important than the quality of the generated cell image. Hence, the researchers not only remove full connecting, but also simplify the structure of the decoder network which does not adopt inception architecture. These methods cut off the number of parameters, although the quality of the generated images is reduced.

The decoder network contains 6-layer up-sample networks, each of which consists of 3 × 3 deconvolutions [46] and ReLU activation. In particularly, the sigmoid function is used to replace ReLU activation of the last layer network for generating images. Here, the designed generator is described in Table 1.

2) JOINT LOSS FUNCTION

The Cell-GAN is trained by comparing the difference between the annotated single-cell images and the corresponding generated single-cell images. In addition to the adversarial loss function, the joint loss function of the Cell-GAN also employs the L_2 loss function [28].

GANs apply a game-like mechanism to train the generator and discriminator. The learning process is that the discriminator D tries to distinguish the real samples x with the generated images, which are produced by the random noise z from G, while G attempts to deceive D by generating samples more similar to x [23].The original adversarial loss function can be expressed as:

$$\min_{G}\max_{D} E_{x\in X}[\log(D(x))] + E_{z\in Z}[\log(1 - D(G(z)))] \quad (1)$$

In proposed approach, the input of generator is no longer a random noise, but a cell image x and its guide factor o. D is used to measure the integrity of the annotated single-cell image $\hat{x}$ and the generated single-cell image. Therefore,

the adversarial loss function L_{adv} can be defined as:

$$L_{\text{adv}} = \min_{G}\max_{D} E_{\hat{x}\in\hat{X}}[\log(D(\hat{x}))] + E_{x\in X}[\log(1 - D(G(x \mid o)))] \quad (2)$$

The L_2 loss function is chosen as part of the joint loss which focuses more on capturing the overall structure of image [26]. Furthermore, divide the generated image into existent region and non-existent region of a cell based on a binary mask M constructed by the annotated cell image, and then the L_2 loss function L_{smi}, can be described by:

$$L_{\text{smi}} = \alpha\|M \cdot (X - G(x \mid o))\|_2 + \beta\|(1 - M) \cdot (X - G(x \mid o))\|_2 \quad (3)$$

α and β are used to adjust the generating scale of Cell-GAN for uncertain cell area, retained or discarded. Specifically, α makes Cell-GAN tend to retain the uncertain cell area as part of the generated cell, which may lead to the contour of generated cell larger than the ground truth, while β is the opposite. Finally, the construction of joint loss functions L_{tot} can format as follows:

$$L_{\text{tot}} = \gamma_{\text{smi}} L_{\text{smi}} + \gamma_{\text{adv}} L_{\text{adv}} \quad (4)$$

and the Adam optimizer is used to train the generator and the discriminator respectively [24].

B. EXTRACTION OF REAL SEGMENTED IMAGE

It is convenient to extract the cell contour from the generated cell image. When the background of the image of the cell to be segmented is complicated, such as highly overlapping, the noise interference is unavoidable in the generated result. The main task of segmentation contour extraction is to denoise the image. Thus, before contour extraction, the guide factor is used to locate the segmentation contour and discard the remaining contours. Then, set the value of binarization threshold to 245, which means that all image information is preserved, except white.

It should be noted that the extraction process is not involved in the training process of Cell-GAN. The researchers tried to integrate them together and three methods were tested: the first method output binary contours of segmented cells, the other two methods output segmented cell images by combining the last layer of decoder network with the input cell image by additive operation or dot product. However, all of them caused the difficulty of convergence of Cell-GAN.

C. IMAGE PREPROCESSING

1) GUIDE FACTOR

The guide factor is used to help Cell-GAN to locate the cell. It is the key to segment the overlapping cells. Without guide factor, multi-cell images would make Cell-GAN fell puzzled about which cell should be segmented, resulting in that a multi-cell image is generated, the contours of which are still inseparable. The researchers choose the nucleus as the guide factor because it is the best mark of the cell and is easy to be detected [47].

In this paper, the guide factor is determined by the approach described in literature [48] based on Depth Information. It should be emphasized that the guide factor is only used for localization, thus the construction of guide factor for nucleus segmentation does not actually require precise segmentation, which reduces the difficulty of the nucleus segmentation task. Figure 3 shows the influence of guide factor on segmentation results. In addition, when the nuclei of two cells are too close to each other, using only the nucleus as the guide factor still makes the segmentation confusing, therefore, part of the cytoplasm around the nucleus can be included in the guide factor to increase the confidence of cell location. At this point, further increasing the amount of information contained in the guide factor does not bring significant improvement to the segmentation performance. This is because the performance of segmentation comes more from the structure of the Cell-GAN network itself rather than from the components of the guide factor. Therefore, the segmentation of cell nuclei only needs to achieve precise localization but not precise segmentation.

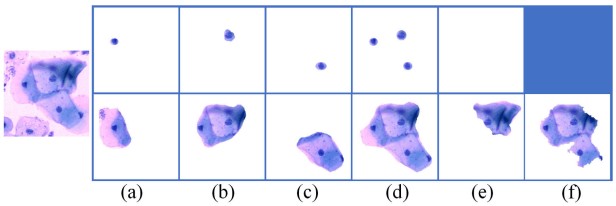


FIGURE 3. Influence of guide factors on segmentation results. From up to down row: guide factors, segmented images. The cell image is cropped by floating box. (a) to (e) are the results of Cell-GAN trained with guide factors. (d) is the result of guide factor containing all of nuclei. (e) is the result of guide factor without any nucleus. (f) is from simplified Cell-GAN trained without guide factor.

2) RECURRENT IMAGE CROPPING

Since the computation of large size images for Cell-GAN is heavy, image cropping is necessary. However, it is difficult to choose a suitable box size to crop cell images. To obtain the appropriate size of boxes, recurrent image cropping (R-crop) is proposed based on the trained Cell-GAN.

The core of R-crop is that for the cell to be segmented, firstly, the approximate contour is gained by use of segmentation result of Cell-GAN, and then use the approximate size to crop the image and re-input it into Cell-GAN. As the iteration progresses, the approximate size becomes closer to the true size of the cell. Finally, when the area change of the segmentation result is less than a certain threshold, the loop is terminated.

For the cell image with high density and large size, the buildup of cells after image resize will cause that the size of segmentation result is close to the input image. Therefore, for R-crop, the first step is coercive crop. In addition, this step helps to reduce the number of iterations. Based on the location of the nucleus, every input image is cropped into a set of images with size of 400 × 400.

The researchers determine whether the segmentation is completed by the change rate of per unit area of the cell I_t for each iteration,

$$I_t = \frac{2 \times [T_t(G(x \mid o)) - T_{t-1}(G(x \mid o))]}{[T_t(G(x \mid o)) + T_{t-1}(G(x \mid o))]} \tag{5}$$

where $T_t(G(x \mid o))$ is the area of the segmentation result. If both I_t and I_{t-1} are less than a specified value, the segmentation is completed,

$$|I_{t-1}| < \rho, \text{ and } |I_{t-1}| < \rho \tag{6}$$

where ρ is the area threshold.

The size of the crop box L is obtained by the sides j and l of the external rectangle for the segmentation result. However, there is a problem that needs to be noticed: The crop size is equal to or smaller than cell size. In the early iterations, especially for large-sized images, the size of the segmentation result might be close to the input image. Furthermore, if the cropping size is smaller than the size of the cell, the cropping collapses into the nucleus. To solve these problems, the dimension change operator ϑ is used to measure the stability of the iteration. ϑ can be defined by:

$$\vartheta_t = \frac{\bar{l}_t}{L_{t-1}} \tag{7}$$

In which, $\bar{l}$ is the longest side of the external rectangle:

$$\bar{l}_t = \max(l_t, j_t) \tag{8}$$

ϑ represents the proportion of the segmentation result in the cropped image. The larger ϑ is, the greater risk of collapse occurs, or it means that the last cropping has caused collapse. When the iterative collapse happens, the segmentation result becomes irregular, so the area ratio is unsuitable for ϑ. Then, according to different ϑ, different cropping rules are defined:

$$L_t = \begin{cases} (1+\varphi) \times \sqrt{l_t^2 + j_t^2} & \vartheta_t < \omega \\ \bar{l}_t + 10 & \vartheta_t > \omega, t = 1 \\ L_{t-1} + \dfrac{(L_{t-2} - L_{t-1})^2}{L_{t-2}} & \vartheta_t > \omega, t > 1 \end{cases} \tag{9}$$

$$L_t = \max(L_t, \delta) \tag{10}$$

where φ is the cropping coefficient, ω is the measure threshold, and δ is the cropping threshold. For Eq. (9), the researchers use the diagonal edges of the directly circumscribed rectangle of the cell outline to obtain the crop frame size to ensure that the cropped image can contain the entire cell, so that the risk of collapse is reduced [49]. However, the diagonal side is easy to result in equal size for early iteration. Hence, the researchers use the longest side to deal with early iteration, and add a quantitative buffer to prevent collapse. For Eq. (10), it limits the minimum size of cropping. In general, it is equal to the input of Cell-GAN meaning that the image is not resized. It helps to prevent iterative collapse.

D. EVALUATION MEASURES

The dice coefficient is one of the most popular measures to evaluate the segmentation quality [51]. Let A denote the region of the automatically segmented shape and let B represent region of the ground truth shape. The dice coefficient(DC) measures the overlapping degree between the two regions and is given by

$$DC = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{11}$$

where $DC \in [0, 1]$. Generally, when DC value is above 0.7, it is considered as an effective segmentation, which means that there is at least 70% similarity between the segmented region and its ground truth. Where, the true positive (TP) area is consisted of pixels that are true in both the segmentation mask and the ground truth mask, while the true negative (TN) area is consisted of pixels that are in neither the segmentation mask and nor the ground truth mask. Similarly, the false negative (FN) area is consisted of pixels that are false in the segmentation mask but true in the ground truth mask, and the false positive (FP) area is consisted of pixels that are in the segmentation mask but not in the ground truth mask. In this paper, we show a pixel-level evaluation using the true positive rate (TPR) and false positive rate (FPR) for evaluating the accuracy of the segmentation.

$$TPR = \frac{TP}{TP + FN} \tag{12}$$

$$FPR = \frac{FP}{TN + FP} \tag{13}$$

In addition, the object-level false negative rate (FNR_o) is significant for computing the underreporting rate of segmentation from the proportion of cells with DC values below 0.7. Here, the object-level true positive (TP_o) term represents the number of cells correctly segmented. The object-level false positive (FP_o) term represents the number of cells incorrectly segmented. The object-level false negative (FN_o) term represents the number of the segmented cells which are really the background. Thus, the object-level false negative rate (FNR_o) can be calculated by (14).

$$FNR_o = \frac{FN_o}{TP_o + FN_o} \tag{14}$$

IV. EXPERIMENT AND ANALYSIS

In this section, the details of the dataset and experiments performed are presented, along with the specific settings, results and analysis.

A. DATASET

The dataset is collected from Harbin Maria Obstetrics and Gynecology Hospital. An automatic pathological scanner was used to acquire images, as show in Figure 4. This equipment consists of an automatic optical microscope, an industrial camera, a personal computer. The computer is equipped with 8G memory, i5-4590 CPU, 3.3GHz.

① Industrial camera: CMOS camera with the resolution of 2048 × 2048 pixels, frame rate 50 frame/s, gray level of 256;
② Camera-microscope interface;
③ Electric platform: walking accuracy for both the X axis and the Y axis is less than 5 μm^2;
④ Remote control bar;
⑤ Microscope: objective lens of 20 times magnification;
⑥ Microscope light box;
⑦ Control box.

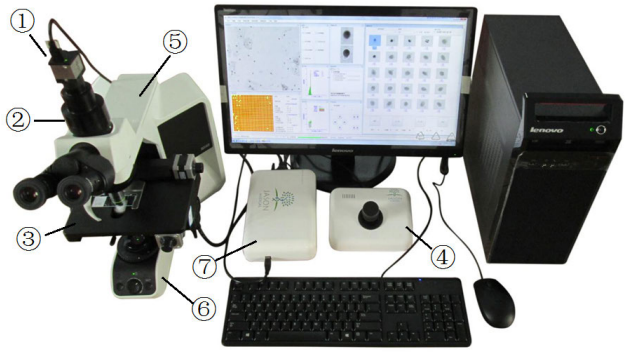


FIGURE 4. Equipment used to acquire images.

This equipment can control the electric platform to automatically move, focus and control the camera to grab images. The process of grab image is set to start from the center of the slide and automatically move like snake, the overlapping pixels is 30 and each slide shot 400 pictures.

In this paper, 60 cervical cell images of size 2048 × 2048 pixels were selected from 12 pathology slides, 40 of which were used to construct the training set and the remaining 20 were used as the test set. After recurrent image cropping, cell nuclei segmentation [48] and manual annotation, each of training and test data is consists of three images: small cell images, guide factors and annotated cell images. Small cell images are the cropped cell images from the original image set. Guide factors are the annotated nuclei used to locate cell to be segmented. Annotated cell images are the annotated complete single-cell images. The details of the dataset is shown in Table 2. The training set consist of 1571 cell images and 520 of which are repetitive but contain different positional information. Some examples of the training data are shown in Figure 5. Each cell is annotated in the cropped image. On the one hand, it brings data enhancement. On the other hand, it helps to strengthen the positioning ability of guide factors. Under different guide factors, the Cell-GAN needs to generate different cell images for the same input image. The test set consists of 309 cell images, including 100 single-cell images, 170 overlapping cell images, and the rest are atypical cell images and test images for ideal cropping.This dataset has been uploaded to the website: https://ieee-dataport.org/documents/cervical-cell-images.

TABLE 2. Details of the dataset.

Category	Original image	Experimental image	Number
The training set	40	single-cell images	1051
		overlapping cell images	520
The test set	20	single-cell images	100
		overlapping cell images	170
		atypical cell images	39

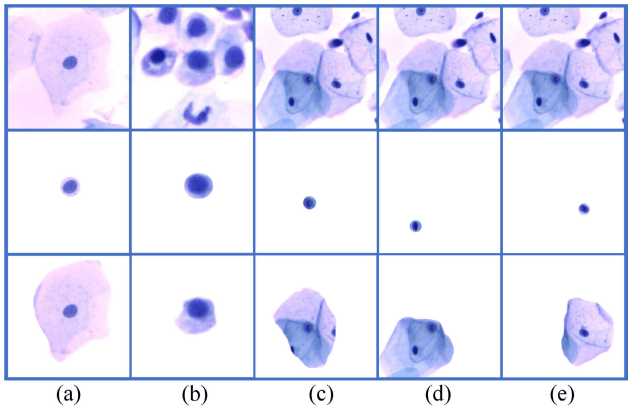


FIGURE 5. Overview of the training set. From up to down row: cropped images, guide factors, annotated images. (a) and (b) are cropped by fixed box. (c) is cropped by floating box. (d) and (e) are annotated data for each cell.

B. EXPERIMENT CONFIGURE

The proposed method use TensorFlow to build Cell-GAN, and then train it by GTX 1080Ti. For the latest version of Cell-GAN, the learning rate, Adam Optimizer momentum and training iteration number are set as 5×10^{-4}, 0.5 and 2×10^3, respectively. In addition, each training image is rotated 4 times to increase the number of samples. For the loss function L_{smi}, two weights are both set to 1, meaning no bias. To avoid overfitting, the output of each layer network is normalized and the dropout with a probability of 0.5 is added to the generator [51]. Finally, the number of parameters in the proposed network is 7767k, around 31.7M.

For R-crop, the maximum number of iterations, the area threshold ρ, the measure threshold ω, the cropping coefficient φ, and the cropping threshold δ are set as 15, 0.15, 0.69, 0.2, and 200, respectively.

For the latest version of non-optimized Cell-GAN, it took us 36 hours to train. After training, the average segmentation speed for cropped cell images is 0.10 second per image.

C. EXPERIMENTAL ANALYSIS

1) INFLUENCE OF GUIDE FACTOR

The guide factor is used to help Cell-GAN to locate the cell. Each cell in the cropped images is annotated and used to train Cell-GAN (see Figure 5). These annotated cells make the Cell-GAN generate different single-cell images when the input images are same but guide factors are different, as shown in Figure 3.

To establish the connection between the guide factor and the cell image in the Cell-GAN, the researchers use additive operation instead of parallel operation to merge the

feature maps. Because experiment shows that parallel operation makes the guide factor lose positioning ability. The conditional GAN [52] uses parallel operation and full matrix which is constructed by image label to control the generated results. By compare the difference between the guide factor and the full matrix, the researchers find that because of the low percentage of the guide factor, such as nucleus, in the whole image, the guide factor is neglected. On the contrary, additive operation plays an important role in strengthening the local information, which makes the weights of the information in some regions more significant than other regions. Figure 6 shows the effect of the guide factor on the generator when the feature maps use additive operation. As a result, the guide factor makes the color of the nucleus in the middle cell different from other nuclei. And these feature maps describe the contours of cells form different views, which is more intuitive after the first additive operation (see Figure 6 (a)).

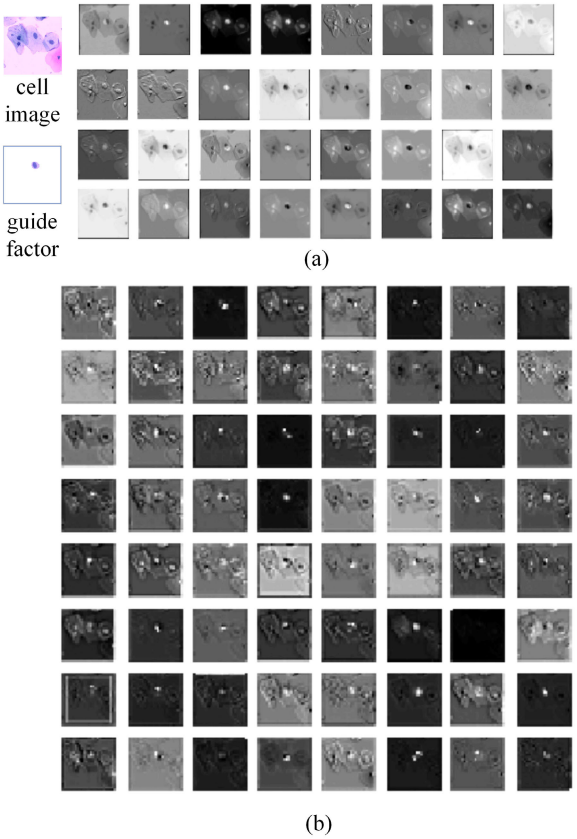


FIGURE 6. Visualization of feature maps after additive operation between cell image and guide factor. The guide factor lights up the nucleus of the middle cell.

In addition, it should be noted that the amount of information contained in the guide factor affects the final segmentation result slightly. It can improve the performance when the segmentation is successful, but it cannot turn over the segmentation failure, even if the ground truth is used as the guide factor, as shown in Figure 7.

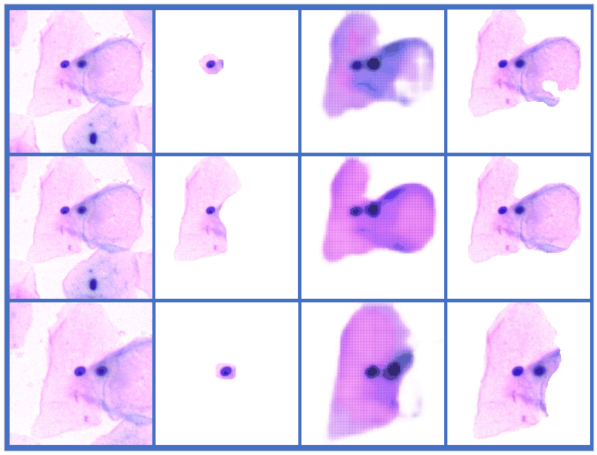


FIGURE 7. Samples of segmentation failure completely. From left to right column: cell images, guide factors, generated results, segmented results. The third row shows the segmentation results under ideal cropping.

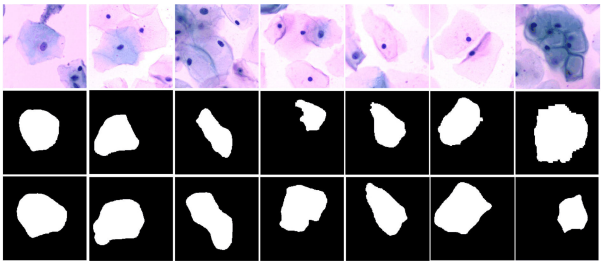


FIGURE 8. Comparison of the effect of different weights of loss functions L_smi on segmentation results. From up to down row: input images, segmentation results when α is 0.1 and β is 10, segmentation results when α is 1 and β is 1. The last column of images represents failed segmentation.

2) INFLUENCE OF DIFFERENT WEIGHTS OF L_{smi}

The effect of different weights of the cost function L_{smi} on the segmentation results is tested. The researchers test two cases: one case is α and β both are equal to 1, and the other is that α is 0.1 and β is 10. The experiment shows that the contours of segmented cells in the first case are larger than the contours in the second case (see Figure 8). During the training process, a large β gives Cell-GAN severe punishment when Cell-GAN generates pixels in non-existent region of cell, while a small α gives it a low reward when segmentation is correct. Therefore, Cell-GAN tends to discard the uncertain regions. But when a cell is in a cell cluster, the Cell-GAN may tend to retain these regions and even expand, because these cells have similar forms so that they are treated as one cell. In addition, it needs to be noted that the changes of contours of segmented cell images in all direction are inconsistent, which is different from erosion and dilation [53]. Hence, the weights adjustment should be carefully.

3) INFLUENCE OF DIFFERENT ITERATIONS OF R-CROP

Figure 9 shows the segmentation performance of Cell-GAN with R-crop at different iterations. The values of ρ, φ, and ω are fixed. At the early iteration, due to the large size of the cropped image, there is under-segmentation in multi-cell overlapping area. The evaluation of iteration termination

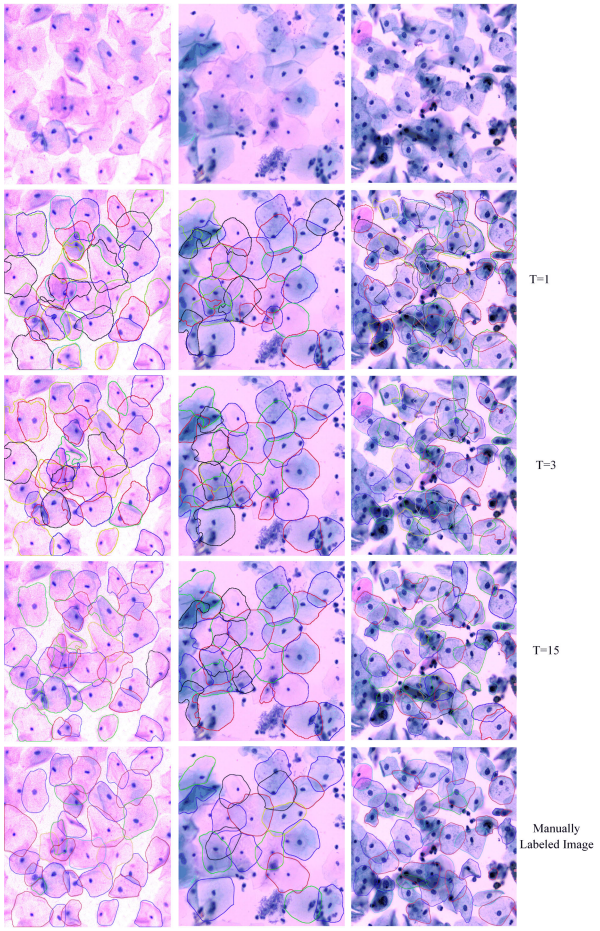


FIGURE 9. Segmentation results of Cell-GAN. Cells close to the boundary are ignored.

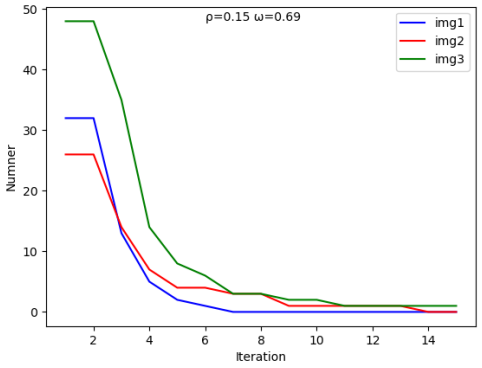


FIGURE 10. Relationship between the iteration number and the Number of remaining cells. From the third iteration, the segmentation termination condition is judged and most cells are segmented in the middle of iteration. Only a small number of cells reach the late iteration, most of which are small cells.

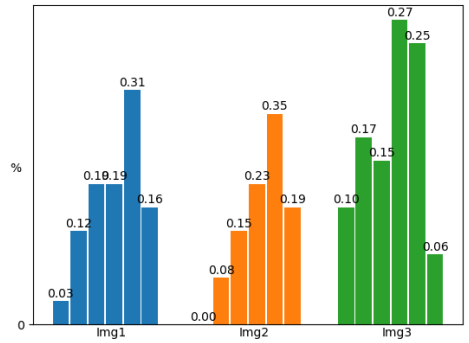


FIGURE 11. Size distribution of final cropped images. For each image, the leftmost column represents 200, the rightmost column represents 400. With 50 pixels as a dividing scale, the remaining sizes are divided into four cases.

begins in the third iteration. Figure 10 shows the change in the number of remaining cropped images. After third iteration, most cropped images have been segmented. However, it does not mean that all cell images can be segmented well. There is a cell in the third image cannot be segmented and makes R-crop into a dead loop. This is because that for a small cell, a tiny change is easy to make the value of I_t bigger than the area threshold ρ. Figure 11 shows the size distribution of final cropped images. Cropped images with the largest and the smallest sizes are regard as two special cases. Then, with 50 pixels as a dividing scale, the remaining cropped images are classified as four cases. A good cropping rule should make the distribution wider and does not contain many images with the largest and the smallest sizes.

4) INFLUENCE OF DIFFERENT PARAMETERS

For proposed segmentation method, there are six parameters to be considered. α and β as essential parameters have been introduced in 3.3.2. The area threshold ρ determines the condition of the iteration termination, is the key of R-crop. Figure 12 shows the change of I_t for several cropped images. Several cropped images are tracked, including the special image that makes R-crop into a dead loop. For most cropped images, the area change tends to be in a regular and small fluctuation. The regular fluctuation means that the segmentation result changes regularly with the number of iterations. The small fluctuation means that the difference between segmentation results is not obvious. As the special case mentioned above, large fluctuation is caused by inaccurate cropping. The cropping threshold δ is used to prevent the collapse, but it also stops R-crop cropping smaller images for small cells. In the experiment, the researchers find that reducing the value of δ can handle the special dead loop. Precise cropping reduces the fluctuation. However, the risk of collapse is also increased. Hence, keeping the limitation of δ is more suitable. The measure threshold φ is the most important parameter affecting the segmentation performance. In the experiment, one can find that large φ is easy to result in equal size, while small φ results in a collapse.

5) INFLUENCE OF CELL OVERLAPPING

In general, the constraint of single-cell generation ensures that Cell-GAN does not care about whether the cells are overlapping or not and treats them in the same way. Figure 13 shows the segmentation results for overlapping cell images cropped by floating box. In general, the ideal cropped

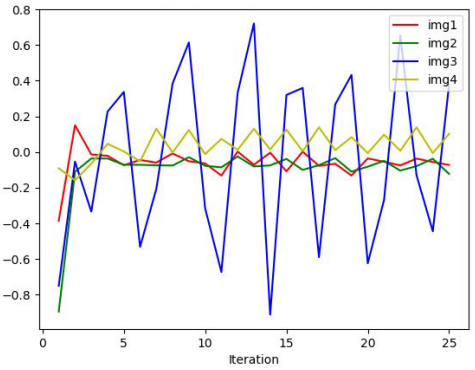


FIGURE 12. The change of I_t for cropped images. Large fluctuations often occur in small cells.

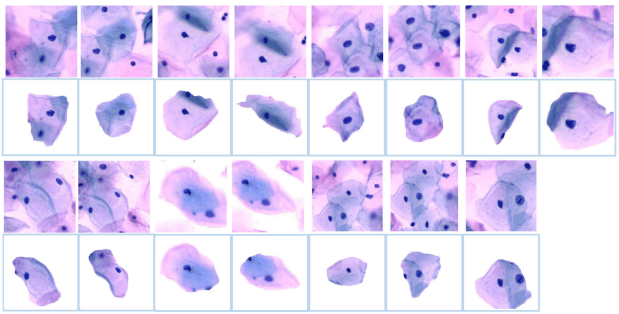


FIGURE 13. Samples of overlapping cells segmentation. Due to the floating box cropping, overlapping cells do not usually appear in one image. The last group of images show the segmentation results for three overlapping cells.

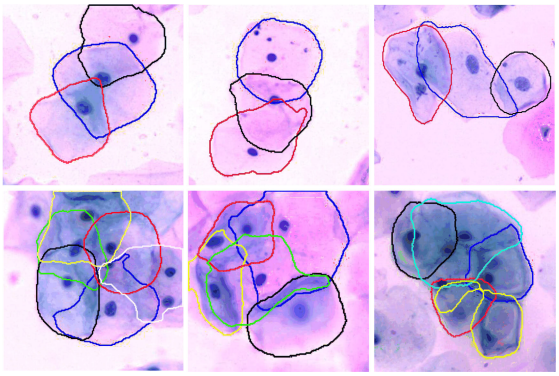


FIGURE 14. Samples of multi-cell images segmentation. There is segmentation failure in the last image. The solution is discussed in Section 4.3.4.

images contain only one complete cell, thus the segmentation results of overlapping cells appear in the different cropped images. However, ideal cropping is hard to be achieved. Especially, due to under-segmentation, the cropped images often contain multiple cells. Therefore, guide factors are introduced to cope with this problem. Moreover, the proposed method is tested by multi-cell images, as shown in Figure 14. To demonstrate the segmentation performance and the positioning ability of guide factors, each cell in the multi-cell images is segmented. As a result, because the positioning ability of guide factors has been enhanced in the training process,

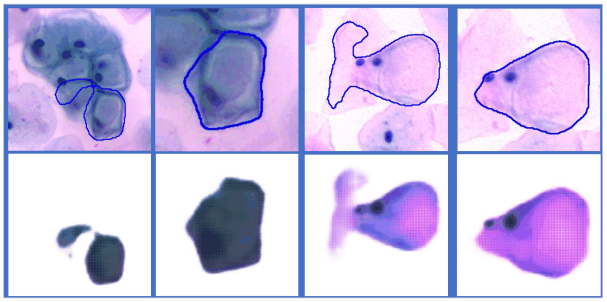


FIGURE 15. Samples of segmentation performance improving. From up to down row: segmented images, corresponding generated images.

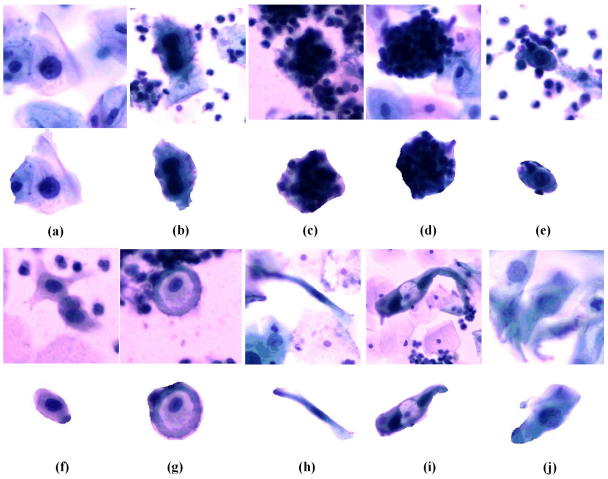


FIGURE 16. Samples of non-encompassed cells segmentation. (a), (b) are cervical cancer. (c), (d) are glandular clusters. (e) contains large number of inflammatory cells. (f) to (j) are non-typical cells.

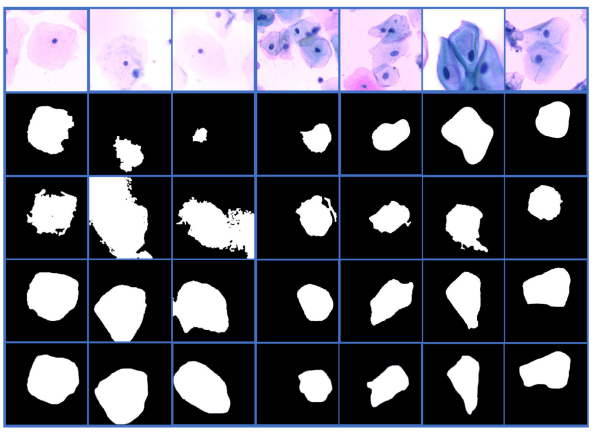


FIGURE 17. Segmentation results for different methods. From up to down row: input images, outlined by Phoulady, Z. Lu, the proposed method and the ground truth.

the proposed method also works well when overlapping degree is high. Furthermore, for serious overlapping that the contours of cells are difficult to be distinguished by human, the segmentation results are acceptable. In the last image, there is segmentation failure, but it can be solved by more suitable cropping and amendment of extraction (see Figure 15).

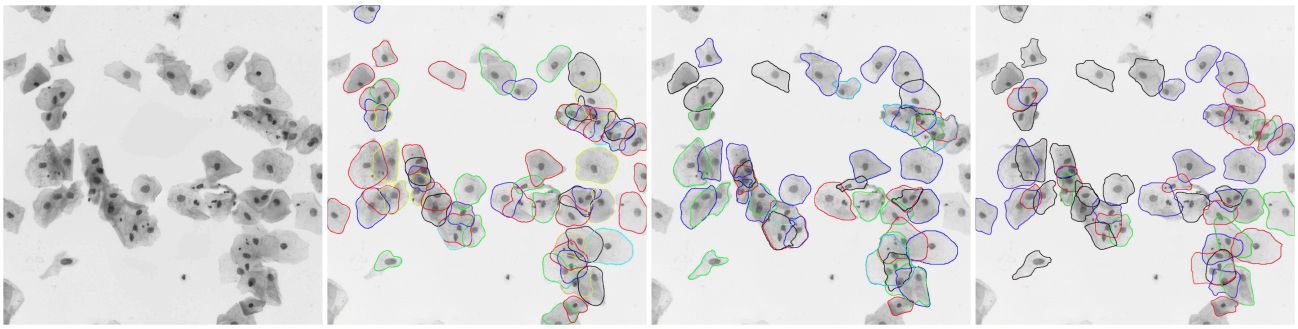


FIGURE 18. Segmentation results on ISBI2015. Cells close to the boundary are ignored for the proposed method.

TABLE 3. Results for various methods of cell region segmentation.

	Single Cell				Overlapping Cell			
	DC	TPR	FPR	FNR_o	DC	TPR	FPR	FNR_o
EMSEG [15]	0.856	0.8	0.008	0.323	0.81	0.723	0.006	0.434
MSSEG [11]	0.864	0.891	0.018	0.287	0.857	0.853	0.036	0.27
Superpixel+CNN [18]	0.886	0.902	0.028	0.256	-	-	-	-
Context Encoders [26]	0.901	0.897	0.011	0.129	0.869	0.885	0.028	0.172
Proposed method	**0.943**	0.979	0.017	**0.079**	**0.899**	0.954	0.031	**0.064**

6) THE GENERALITY OF CELL-GAN

Because the training set is not classified, it includes three types of cells: the upper squamous cells, the mid-level squamous cells, and the basal cells. The researchers test the segmentation performance for cervical cancer cells, glandular clusters, and special-shaped cells, as shown in Figure 16. The net can segment these special cells effectively even if they are not used to train Cell-GAN. The glandular clusters have different forms from other cells. They are tested as a basis for glandular cancer detection which needs further study.

V. EVALUATION

After reviewing the literatures on cervical cell segmentation in recent years, it was found that this work mainly focused on nuclear segmentation and few methods for cytoplasmic segmentation [54]. The CELL-GAN proposed in this paper can achieve a good effect of cytoplasmic segmentation, here choose Phoulady's method EMSEG [15] and Lu's method MSSEG [11] for comparison. The former is based on the EM algorithm, while the latter is based on multiple level sets. Both of them are evaluated on ISBI2015 using traditional methods and have excellent performance. The results of the three compared methods on own database are shown in Figure 17. It shows that when the contrast of images is poor and cells are highly overlapping, the cell contour information is difficult to obtain, however, in these cases, the proposed method can achieve better performance than that obtained by Phoulady's and Lu's methods.

The researchers test the Cell-GAN on the ISBI2015 database, as shown in Figure 18. It should be noted that the Cell-GAN is not fine-tuned by the ISBI2015 training set and does not use the stack of multi-focal images. All parameters are set as default values, except for the cropping threshold δ. Where $\delta = 150$, because cell number in ISBI2015 database is smaller than private database. The default parameter α and β are both set to 1, the area change threshold ρ is set to 0.15, the scale change threshold ω is set to 0.69, the cropping buffer coefficient φ is set to 0.2, and the minimum cut size δ is set to 200. The maximum number of iterations is set to 15.

Further, excellent deep-learning-based cell segmentation methods are chosen for comparison. Song Y *et al.* [18] propose the segmentation method based on CNN and D. Pathak *et al.* [26] propose Context Encoders which is purposed to repair images based on GAN, the researchers made it available for segmentation of cell images by adding guide factors to demonstrate the advantage of Cell-GAN. Given that Song Y's method cannot segment overlapping cells only by using CNNs, the performance of these methods for single-cell images and overlapping cell images are evaluated respectively. As a result, the proposed method achieves the best performance both on DC and FNR_o, the detail is shown in Table 3. Because the single-cell images are elected with poor contrast, it causes that the value of FNR_o for single-cell images is bigger than overlapping cell images. In addition, serious overlapping is not included because annotation is hard.

VI. CONCLUSION

In this paper, a new approach is proposed to segment cervical cells based on the Cell-GAN. For each cell to be segmented in the cell image, the Cell-GAN generates a single-cell image by judging the integrity of a cell based

on the probability distribution of cell morphology. And the contour of the generated cell serves the final segmentation line. In the process of segmentation, the guide factor used to locate the cell guarantees the success of segmentation, especially for overlapping cells. As an automatic image cropping method, R-crop further improves the segmentation performance.Overall, the proposed method well grasps the common properties of cervical cells by the estimation of probability distribution and effectively deals with the cell-overlapping problems. Moreover, the guide factor and the R-crop operation further improve the segmentation accuracy and computational efficiency. The segmentation results show that the proposed Cell-GAN method can achieve a comprehensive segmentation capability, especially in the case of highly overlapping cells and complex backgrounds, the segmentation failure rate is significantly lower than that of other four methods.

The proposed method may be applied to some kinds of histopathology image cell segmentation, such as prostate and breast cancer histopathology images. Another potential application may be blood image cell segmentation. However, there are still some issues that need further study in the proposed method, such as how to improve the performance of R-crop and how to segment cells that have adjacent nuclei. In the future, the researchers also need to improve the applicability of proposed method so that it can meet the segmentation of more categories of cells at the same time.

REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clinicians*, vol. 68, no. 6, pp. 394–424, Nov. 2018, doi: 10.3322/caac.21492.

[2] M. Arbyn, E. Weiderpass, L. Bruni, S. de Sanjosé, M. Saraiya, J. Ferlay, and F. Bray, "Estimates of incidence and mortality of cervical cancer in 2018: A worldwide analysis," *Lancet Global Health*, vol. 8, no. 2, pp. e191–e203, 2020, doi: 10.1016/S2214-109X(19)30482-6.

[3] A. Ray, I. K. Maitra, and D. Bhattacharyya, "Detection of cervical cancer at an early stage using hybrid segmentation techniques from PAP smear images," *Int. J. Adv. Sci. Technol.*, vol. 112, pp. 23–32, Mar. 2018, doi: 10.14257/ijast.2018.112.03.

[4] G. Naizhaer, J. Yuan, P. Mijiti, K. Aierken, G. Abulizi, and Y. Qiao, "Evaluation of multiple screening methods for cervical cancers in rural areas of Xinjiang, China," *Medicine*, vol. 99, no. 6, 2020, Art. no. e19135, doi: 10.1097/MD.0000000000019135.

[5] A. Sarwar, A. A. Sheikh, J. Manhas, and V. Sharma, "Segmentation of cervical cells for automated screening of cervical cancer: A review," *Artif. Intell. Rev.*, vol. 53, no. 4, pp. 2341–2379, Apr. 2020, doi: 10.1007/s10462-019-09735-2.

[6] S. S. Geetha, "Enhancing the classification of pap smear images using ENN–TLBO classification method," *Int. J. Innov. Technol. Exploring Eng.*, vol. 9, no. 6, pp. 553–558, 2020, doi: 10.35940/ijitee.f3708.049620.

[7] Y. Al-Kofahi, A. Zaltsman, R. Graves, W. Marshall, and M. Rusu, "A deep learning-based algorithm for 2-D cell segmentation in microscopy images," *BMC Bioinf.*, vol. 19, no. 1, pp. 1–11, Dec. 2018, doi: 10.1186/s12859-018-2375-z.

[8] J. Ao, S. Mitra, R. Long, B. Nutter, and S. Antani, "A hybrid watershed method for cell image segmentation," in *Proc. IEEE Southwest Symp. Image Anal. Interpretation*, Apr. 2012, pp. 29–32, doi: 10.1109/SSIAI.2012.6202445.

[9] S. S. Pfister, M. Betizeau, C. Dehay, and R. Douglas, "Robust 3D cell segmentation by local region growing in convex volumes," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, Apr. 2013, pp. 426–431, doi: 10.1109/ISBI.2013.6556503.

[10] S. Kaur and J. Sahambi, "Curvelet initialized level set cell segmentation for touching cells in low contrast images," *Comput. Med. Imag. Graph.*, vol. 49, pp. 46–57, Apr. 2016, doi: 10.1016/j.compmedimag.2016.01.002.

[11] Z. Lu, G. Carneiro, and A. Bradley, "An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1261–1272, Apr. 2015, doi: 10.1109/TIP.2015.2389619.

[12] N. Mat-Isa, M. Mashor, and N. Othman, "Seeded region growing features extraction algorithm; its potential use in improving screening for cervical cancer," *Int. J. Comput. Internet Manag.*, vol. 13, no. 1, pp. 61–70, 2005.

[13] T. R. Jones, A. Carpenter, and P. Golland, "Voronoi-based segmentation of cells on image manifolds," in *Proc. Int. Workshop Comput. Vis. Biomed. Image Appl.*, vol. 3765. Berlin, Germany: Springer, 2005, pp. 535–543.

[14] A. Kale and S. Aksoy, "Segmentation of cervical cell images," in *Proc. 20th Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 2399–2402, doi: 10.1109/ICPR.2010.587.

[15] H. A. Phoulady, D. B. Goldgof, L. O. Hall, and P. R. Mouton, "A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 201–204, doi: 10.1109/ISBI.2016.7493244.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. 26th Annu. Conf. Neural Inf. Process. Syst.*, Lake Tahoe, NV, USA, vol. 2, Dec. 2012, pp. 1097–1105, doi: 10.1061/(ASCE)GT.1943-5606.0001284.

[17] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017, doi: 10.1109/TPAMI.2016.2572683.

[18] Y. Song, L. Zhang, S. Chen, D. Ni, B. Li, Y. Zhou, B. Lei, and T. Wang, "A deep learning based framework for accurate segmentation of cervical cytoplasm and nuclei," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 2903–2906, doi: 10.1109/EMBC.2014.6944230.

[19] Y. Song, E. L. Tan, X. Jiang, J. Z. Cheng, D. Ni, S. Chen, B. Lei, and T. Wang, "Accurate cervical cell segmentation from overlapping clumps in pap smear images," *IEEE Trans. Med. Imag.*, vol. 36, no. 1, pp. 288–300, Jan. 2017, doi: 10.1109/TMI.2016.2606380.

[20] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351. Cham, Switzerland: Springer, 2015, pp. 234–241.

[21] J. Yu, M. Tan, H. Zhang, D. Tao, and Y. Rui, "Hierarchical deep click feature prediction for fine-grained image recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jul. 30, 2019, doi: 10.1109/tpami.2019.2932058.

[22] G. Ting, W. Weixing, L. Wei, and Y. Dandan, "Rock particle image segmentation based on improved normalized cut," *Int. J. Control Autom.*, vol. 10, no. 4, pp. 271–286, Apr. 2017, doi: 10.14257/ijca.2017.10.4.24.

[23] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, vol. 3, 2014, pp. 2672–2680.

[24] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proc. 4th Int. Conf. Learn. Represent. (ICLR) Conf. Track*, 2016, pp. 1–16.

[25] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015.

[26] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2536–2544, doi: 10.1109/CVPR.2016.278.

[27] W. Wang, Q. Huang, S. You, C. Yang, and U. Neumann, "Shape inpainting using 3D generative adversarial network and recurrent convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2317–2325, doi: 10.1109/ICCV.2017.252.

[28] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 105–114, doi: 10.1109/CVPR.2017.19.

[29] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976, doi: 10.1109/CVPR.2017.632.

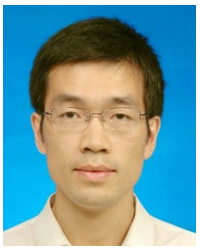
[30] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial text to image synthesis," in *Proc. 33rd Int. Conf. Mach. Learn., (ICML)*, vol. 3, 2016, pp. 1681–1690.

[31] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 2, pp. 84–90, Jun. 2012, doi: 10.1145/3065386.
[32] Y. Tan, *GPU-Based Parallel Implementation of Swarm Intelligence Algorithms*. San Francisco, CA, USA: Elsevier, 2016.
[33] Y. Song, L. Zhang, S. Chen, D. Ni, B. Lei, and T. Wang, "Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 10, pp. 2421–2433, Oct. 2015, doi: 10.1109/TBME.2015.2430895.
[34] A. Tareef, Y. Song, H. Huang, Y. Wang, D. Feng, M. Chen, and W. Cai, "Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling," *Neurocomputing*, vol. 248, pp. 28–40, Jul. 2017, doi: 10.1016/j.neucom.2017.01.093.
[35] Kurnianingsih, K. H. S. Allehaibi, L. E. Nugroho, Widyawan, L. Lazuardi, A. S. Prabuwono, and T. Mantoro, "Segmentation and classification of cervical cells using deep learning," *IEEE Access*, vol. 7, pp. 116925–116941, 2019, doi: 10.1109/ACCESS.2019.2936017.
[36] Y. Zhou, H. Chen, J. Xu, Q. Dou, and P. A. Heng, "IRNet: Instance relation network for overlapping cervical cell segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 11764. Cham, Switzerland: Springer, 2019, pp. 640–648.
[37] N. Souly, C. Spampinato, and M. Shah, "Semi supervised semantic segmentation using generative adversarial network," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5689–5697, doi: 10.1109/ICCV.2017.606.
[38] A. Arbelle and T. R. Raviv, "Microscopy cell segmentation via adversarial neural networks," 2017, *arXiv:1709.05860*. [Online]. Available: http://arxiv.org/abs/1709.05860
[39] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232, doi: 10.1109/ICCV.2017.244.
[40] Y. Huo, Z. Xu, S. Bao, A. Assad, R. G. Abramson, and B. A. Landman, "Adversarial synthesis learning enables segmentation without target modality ground truth," 2017, *arXiv:1712.07695*. [Online]. Available: http://arxiv.org/abs/1712.07695
[41] L. Han and Z. Yin, "Transferring microscopy image modalities with conditional generative adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 99–107, doi: 10.1109/CVPRW.2017.118.
[42] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, 2008, pp. 1096–1103, doi: 10.1145/1390156.1390294.
[43] T. Xiao, Y. Xu, K. Yang, J. Zhang, Y. Peng, and Z. Zhang, "The application of two-level attention models in deep convolutional neural network for fine-grained image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 842–850, doi: 10.1109/CVPR.2015.7298685.
[44] S. A. Tuama and J. H. Saud, "An efficient segmentation method for automated tongue extraction using HSV color model," *Int. J. Adv. Sci. Technol.*, vol. 133, pp. 1–10, Dec. 2019, doi: 10.33832/ijast.2019.133.01.
[45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
[46] L. Xu, J. S. J. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, vol. 2, Dec. 2014, pp. 1790–1798.
[47] H. Lee and J. Kim, "Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and cell-wise contour refinement," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2016, pp. 1367–1373, doi: 10.1109/CVPRW.2016.172.
[48] T. Wang, J. Huang, D. Zheng, and Y. He, "Nucleus segmentation of cervical cytology images based on depth information," *IEEE Access*, vol. 8, pp. 75846–75859, 2020, doi: 10.1109/ACCESS.2020.2989369.
[49] I. . and I. M. Sudarma, "Automatic segmentation of U-Zone area on facial images using fuzzy edge detection," *Int. J. Adv. Sci. Technol.*, vol. 133, pp. 19–30, Dec. 2019, doi: 10.33832/ijast.2019.133.03.
[50] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945, doi: 10.2307/1932409.
[51] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, Jan. 2014.
[52] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*. [Online]. Available: http://arxiv.org/abs/1411.1784
[53] I. Ragnemalm, "Fast erosion and dilation by contour processing and thresholding of distance maps," *Pattern Recognit. Lett.*, vol. 13, no. 3, pp. 161–166, 1992, doi: 10.1016/0167-8655(92)90055-5.
[54] M. M. Rahaman, C. Li, X. Wu, Y. Yao, Z. Hu, T. Jiang, X. Li, and S. Qi, "A survey for cervical cytopathology image analysis using deep learning," *IEEE Access*, vol. 8, pp. 61687–61710, 2020, doi: 10.1109/ACCESS.2020.2983186.

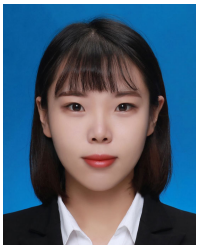
JINJIE HUANG received the B.Sc. and M.Sc. degrees from Harbin University of Science and Technology, in 1990 and 1997, respectively, and the Ph.D. degree from Harbin Institute of Technology, in 2004. He is currently a Professor with Harbin University of Science and Technology. His main research interests include intelligent control of complex industrial process, pattern recognition, and image processing.

GUIHUA YANG received the B.Sc. degree from Northeast Forestry University, in 2007, and the M.Sc. degree from Harbin University of Science and Technology, in 2010, where she is currently pursuing the Ph.D. degree. Her main research interests include medical image segmentation, medical text classification, and disease detection.

BIAO LI received the B.Sc. and M.Sc. degrees from Harbin University of Science and Technology. His main research interests include medical image processing and prediction of diseases in medical images.



YONGJUN HE received the B.S. degree in electrical engineering from Harbin University of Science and Technology, Harbin, China, in 2003, and the M.S. and Ph.D. degrees from the School of Computer Science, Harbin Institute of Technology, Harbin, in 2006 and 2008, respectively. He is currently a Professor with the School of Computer Science and Technology, Harbin University of Science and Technology. His research interests include speech speaker recognition, machine learning, image processing, and speech processing.



YANI LIANG received the bachelor's degree in engineering from Harbin University of Science and Technology, in 2018, where she is currently pursuing the master's degree. Her main research interests include medical image segmentation and medical image classification.