

Research Article

Overlapping Cell Segmentation of Cervical Cytology Images Based on Nuclear Radial Boundary Enhancement

Tao Wang ¹, Hong Lan ¹, Landing Niu ¹, Zhipeng Fan ¹ and Guihua Yang ²

¹School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028, China

²Daqing Normal University, Daqing 163311, China

Correspondence should be addressed to Landing Niu; n_nld@163.com

Received 20 January 2022; Revised 3 March 2022; Accepted 28 June 2022; Published 20 July 2022

Academic Editor: Hye-jin Kim

Copyright © 2022 Tao Wang 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.

The accurate segmentation of cervical cell images is one of the key steps of the cervical cancer computer-aided diagnosis system. For the problem of overlapping cell and boundary blurring in cervical cell clusters, the researchers propose a segmentation algorithm based on the nuclear radial boundary enhancement for overlapping cell of cervical cytology images. This method not only suppresses the noise of cervical cytology images but also preserves the contrast of overlapping cell boundary. The researchers generate the weight graph by the candidate contour points and contour line segment attributes and utilize the dynamic programming algorithm to find the shortest path in the weight graph. The shortest path corresponds to the coarse segmentation contour in the cell image. The level set model is used to finely segment the obtained coarse cell segmentation boundary, so as to obtain the final cervical cell boundary. Through the quantitative and qualitative evaluation results, such as dice similarity coefficient, true positive rate, and false positive rate, it can be seen that the overlapping cell segmentation algorithm in this paper has achieved better segmentation results. Compared with other current overlap cell segmentation algorithms, the segmentation results obtained in this paper have greater advantages.

1. Introduction

The accurate segmentation of cervical cell images is one of the key steps of the cervical cancer computer-aided diagnosis system. After years of development, segmentation technology of the cervical cytology images has made vast improvements and the performance of the cervical cancer computer-aided diagnosis system has been greatly improved. However, due to the characteristics of the actual cervical cytology images and the complexity of the cell morphology, the current segmentation technology still needs to be improved.

Cytoplasmic characteristics have been shown to be critical for the identification of abnormal cells [1]. The accurate segmentation of cytoplasm in cervical cytology images is the core part of cervical cell segmentation. Once the cytoplasmic boundary is located, quantitative evaluation index values such as cell diameter and nuclear-to-cytoplasmic ratio can be calculated. Some early methods solved

the problems of segmentation of free cells which did not overlap with neighbouring cells or only a small part of overlap and partially overlapping cells [2–7].

When the intensity of the cervical cytology images meets the bimodal distribution, the automatic threshold segmentation method [3, 5] can get better segmentation results. Wu et al. [5] used threshold technology to identify the cell in single-cell image. In the poor contrast and uneven staining cervical cytology images, the segmentation results obtained by the threshold method were not ideal. Harandi et al. [3] proposed a multi-resolution threshold cell clump and cytoplasm detection method. This method locates cell clump in low-resolution images, and detects cytoplasm and cell nucleus boundaries in high-resolution images. And the geometric activity model is used to segment the cytoplasm of the single-cell cervical image and the cell image with a low overlap rate.

The active contour model based on gradient vector flow-Active Contour Model (GVF-ACM) [8] is not only robust to

the initial contour, but also has the advantage of converging to the cell concave boundary, so this method is widely used in single Cytoplasmic segmentation of cell images. However, the GVF-ACM method cannot obtain better segmentation results under the situation of poor contour contrast. To solve this problem, Changkong et al. [6] proposed a Cytoplasm and Nucleus Contour (CNC) detector for extracting the nuclear and cytoplasmic regions in cervical cytology images. At the same time, this method utilizes the K-means method to segment the cytoplasm from the background area of the cervical cell image. When malignant cells are exposed to cytotoxic effects, Song et al. [9] proposed a process in which cysteamine is accumulated in diseased cells and converted to DMS on a massive scale. This method is not only reliable for identifying malignant cell lines, but it may also be used to investigate how drug-related cytotoxic effects affect cancer cell metabolism.

Yang-Mao et al. [2] proposed an Edge Enhancement Nucleus and Cytoplasm Contour (EENCC) based on the research results of Changkong [6]. EENCC effectively removes the Gaussian noise and impact noise in the cervical cell image by truncating median filter, and retains the sharp boundary. The main contribution of this detector is the Mean Vector Difference (MVD) algorithm, which assigns a directional gradient vector flow to each pixel in the image, and finally detects the nucleus and the cytoplasmic boundary by the Otsu threshold method. However, this method cannot solve the problems of boundary closure and continuity, and the method based on GVF is too sensitive to noise. By adjusting the level of the speckle noise, Bhardwaj et al. [10] showed that the mean filter offers better results than the median filter on the sampled ultrasonic images. This experiment uses extremely simple filtering techniques that can be improved by utilizing a variety of other filters such as bilateral, trilateral, wavelet-based, and entropy-based filters, among others.

Tsai et al. [7] combined unsupervised classification algorithm and ACM model for cytoplasmic segmentation of single-cell images. The ACM model based on boundaries and regions can accurately detect pixel-level closed target boundaries. Therefore, this method solves the problems of boundary closure and continuity in the literature [2]. In order to solve the noise interference problem of the GVF algorithm, Li et al. [4] proposed the radial GVF snack algorithm on the basis of the multi-directional GVF snack algorithm [11]. The algorithm calculates the gradient of the radial gradient value of each pixel in the image on the radial line segment of the specified nucleus, and uses the radial gradient value to replace the gradient in the GVF algorithm. The ability of the GVF method based on radial gradient to locate the fuzzy boundary is significantly improved, and the gradient distribution along the radial direction can reduce the false gradient caused by dyeing and uneven illumination. This method can accurately segment the cytoplasm and nucleus in an image containing only one cervical cell.

Generally speaking, overlapping cell segmentation is the most challenging task in cervical cytology image segmentation. In the process of cell sample preparation, uneven illumination and differences in dye concentration will result

in low contrast between different components in the cell image. Other factors, such as drying cells, red blood cells, mucus, bacteria, and white blood cells, will increase the difficulty of overlapping cell segmentation [12, 13]. In addition, complex cell shapes and a high overlap rate between cells will also decrease the contrast of cell boundaries. In order to promote the development of overlapping cervical cell image segmentation technology [14, 15], the IEEE International Biomedical Imaging Society held the first and second cervical cell image segmentation challenges in 2014 and 2015, respectively. The dataset of the challenge is publicly available, and the dataset contains training set and test set, and the training set gives the results of manual annotation. This makes it possible to evaluate and compare different overlapping cell segmentation methods. Therefore, most of the overlapping cell segmentation methods in recent years are proposed to solve the overlapping cell segmentation task of this dataset.

It was Ushizima team [16] that won the championship for the first time in the challenge. Ushizima et al. used graph-based linear time algorithm [17] and global search truncation algorithm [18] to obtain cell clumps and nuclear regions based on the similarity of the intensity of adjacent pixels. The overlapping cells in the cell clump are divided into convex polygons through the narrow-band seed region of the nucleus, the graph-based region growth algorithm [19] and the Voeonoi diagram method [20], as shown in Figure 1. It can be seen from the figure that this method can only divide the overlapping area of cells by a straight line, and cannot obtain the area between overlapping cells. Therefore, the accuracy of cell segmentation by this method has a great improvement [21].

The champion of the second challenge was Phoulady's team [22], which adopted the iterative threshold method and the regularization level set evolution algorithm based on the ellipse shape assumption. The iterative threshold is used to segment the cell clump and cell nucleus area, and the cell clump area is divided into windows of specified size. According to the attribute values such as the intensity mean and variance of the window, the area closest to the nucleus of the window is detected to complete the segmentation of overlapping cells. Finally, the smooth cytoplasmic boundary is obtained by the evolutionary algorithm of the regular level set assumed by the ellipse shape. On this basis, the literature [23] defines a focus metric for each subimage in the image, and assigns a nucleus to each subimage according to the position information and focus similarity. A two-step method including subimage-level boundary coarse segmentation and pixel-level boundary refinement segmentation is used to complete overlapping cell segmentation.

In order to obtain a more accurate cytoplasmic segmentation boundary, Phoulady et al. [24] improved the two-step boundary segmentation algorithm. This method finds new boundary candidate points by defining a weight vector, and uses a smoothing filter to smooth the candidate boundary points. Identify the outliers in the candidate boundary points by the distance between each candidate point and its new position. The smoothing filter is used again on the candidate boundary points from which outliers are

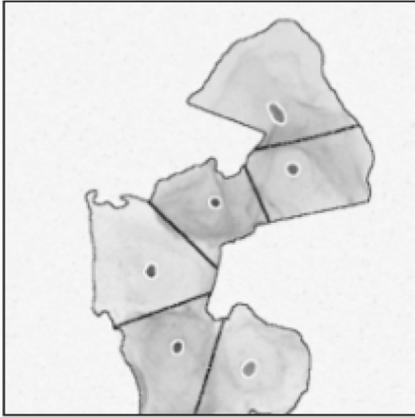


FIGURE 1: Cytoplasmic segmentation based on Voeonoi diagram.

deleted to obtain the final cytoplasmic boundary. The greatest contribution of the literature [24] is the use of the depth information of the stacked images to obtain a more accurate coarse cytoplasmic segmentation boundary.

On the basis of the ellipse shape prior model in literature [22], Nosrati and Hamarneh [25] optimized the algorithm in literature [17] by using the star shape prior to replace the original elliptical shape prior and obtained more accurate cytoplasm segmentation result. These methods prove that combining shape priors into the parameter segmentation process can significantly enhance the segmentation effect of overlapping cells. However, the ellipse shape prior and the star shape prior are too simple to represent the true shape of cervical cells. The overlapping cell segmentation algorithm using a fixed shape prior cannot provide sufficient shape information for the overlapping part of the cell. In order to solve this problem, in literature [26], the authors proposed an automatic shape prior algorithm based on sparse approximation to segment overlapping cervical cells. In the literature [27], Tareef et al. improved the method, by using methods based on deep learning and dynamic shape models to separate individual nuclei and cytoplasm. This method improves the performance of cell nucleus detection and solves the problem of low sensitivity of cervical cell segmentation, and improves the accuracy of overlapping cervical cell segmentation.

In view of the computational efficiency of the super-pixel method, some scholars utilize an improved super-pixel method to complete overlapping cervical cell segmentation. In literature [26], the authors used a fast migration algorithm to generate super-pixel images, and used unsupervised binary classifiers and MSER algorithms to complete cell clump and nuclei segmentation. Finally, the overlapping cytoplasm is segmented by the combination of multi-level set functions. Huang et al. [28] used the SLIC method, adaptive threshold method, and local threshold method to obtain super-pixel maps, cell clusters, and cell nucleus regions, respectively. The method of assigning each super-pixel to the nearest cell nucleus is used to complete the coarse segmentation of the cervical overlapping cell image. Then the fine segmentation of the

coarse segmentation boundary is completed by the graph cut algorithm. Since this method uses the “closest distance” method to segment overlapping cells, this method can only achieve better segmentation results if the nucleus is roughly at the centre of the cell.

With the development of artificial neural network technology, some neural network methods have also been applied to overlap cervical cell segmentation [29]. Song et al. [30] used a segmentation framework based on deep learning and deformation models to segment overlapping cells. This framework not only explores the shape relevance of overlapping cells, but also uses cell structure, background information and multi-cell labelling models to complete overlapping cell segmentation. The multi-label model transforms the segmentation problem into a discrete point labelling problem, and marks the cytoplasmic pixels as corresponding nuclear labels obtained by the CNN method. Then use Gaussian kernel fitting to determine the coarse segmentation boundary of the cell nucleus, and finally use the dynamic multi-template deformation model to refine the segmentation of the cell boundary. CNN has the disadvantages of requiring a large number of samples and high computational complexity.

In cervical cell images, the segmentation of overlapping cell is largely dependent on image quality. As the researchers all know, when the image is polluted by noise, how to complete the denoising operation of the image while retaining the boundary features is a very difficult problem. Typically, cervical cell images are affected by Gaussian noise and shock noise. In the early years, many scholars have carried out extensive research on how to remove noise in images, such as median filtering, Gaussian filtering, and Type-B filtering. Early filtering methods can remove Gaussian noise and shock noise in the image very well, but these methods also reduce the resolution of cell boundaries. Based on this problem, Li et al. [31] proposed an improved median filter (Trim-meaning), which can effectively remove the impact noise and Gaussian noise, and retain the sharp part of the boundary. The method extracts the pixel value in the neighbourhood window of each pixel in the image, arranges it in ascending order, deletes the pixel value of the head and tail according to a certain size, and then replaces the pixel with the mean value. Although this method can retain sharp boundaries, the boundaries of overlapping cells are indistinguishable for cervical cell images [32].

Guan et al. [33] proposed an edge enhancement method based on gradient decomposition. The method first uses morphological filtering to denoise the cell image, and then uses the direction of the gradient value at the pixel point and the angle between the point and the centre of the nucleus to enhance the cytoplasmic boundary belonging to the nucleus. This method needs to calculate the gradient and radial direction angle of each pixel point, which is overly dependent on the condition that the cells need to satisfy the convex boundary. Many boundaries of real overlapping cervical cells do not meet the condition of convex boundary, so this method cannot obtain the ideal boundary enhancement effect.

2. Methods

2.1. Preprocessing. In the preprocessing step, this paper uses the filtering and gradient calculation algorithm based on the radial region of the nucleus to complete the image denoising and gradient calculation tasks. The flowchart of the cervical cell segmentation method in this paper is shown in Figure 2.

2.1.1. Filtering Based on the Radial Region of the Nucleus. In the overlapping clumps of cervical cell images, different cells overlap each other, making their boundaries more blurred. The structural analysis of cervical cells shows that most of the nuclei are located in the cells [34]. Although the positions vary greatly, the cytoplasmic boundary surrounds the nucleus. Cell boundaries exhibit different orientations in different radial directions of the cell. Based on this observation, a filtering algorithm for the radial region of the nucleus is proposed, which can not only effectively remove the noise in this region, but also preserve the resolution of the cytoplasmic boundary to the maximum extent.

C_i is the specified cell, (x_c, y_c) is the centre of the cell, and r_θ is the radial with (x_c, y_c) as the centre and θ as the direction. The coordinate of the point (x, y) on the radial is (θ, ρ) , where ρ is the Euclidean distance between (x, y) and (x_c, y_c) . According to the above definition, with the nucleus as the centre, the cell area is divided into several radial areas according to the step size $\Delta\theta$. For different radial regions, different kernel functions are utilized to denoise the regions. If the step size is too large, the robustness of the algorithm is poor, and if the step size is too small, the corresponding cytoplasmic boundary cannot be preserved. $\Delta\theta = 22.5$ is selected experimentally, and the image was divided into 16 intervals. When $\Delta\theta = 22.5$, the division result of the radial area is shown in Figure 3.

According to the value of $\Delta\theta$, this paper sets four denoising kernel functions and the kernel function settings are shown in Figure 4.

For the pixel point (x, y) on the cell image, select the corresponding kernel function kernel(i) according to its position in the specified cell. The selection of the kernel function is shown as follows:

$$\text{kernel}(i) = \text{floor}\left(\text{mod}\left(\text{ceil}\left(\frac{\theta}{\Delta\theta}\right), \frac{180}{\Delta\theta}\right), 2\right). \quad (1)$$

In the formula, θ represents the radial direction of the pixel; $\text{ceil}(\cdot)$ is the round-up operation; $\text{mod}(\cdot)$ is the remainder operation; $\text{floor}(\cdot)$ is the round-down operation. When the (x, y) point on the image finds the corresponding kernel function kernel(i), the window area $W_{x,y}$ centered on the (x, y) point is convolved with the selected kernel function kernel(i), as shown in

$$b_{x,y} = |\text{kernel}(i) \oplus W_{x,y}|. \quad (2)$$

Here, $b_{x,y}$ is the indicator variable of the contour points in the region, and \oplus is the convolution operation. Comparison of filtering results of different methods is shown in Figure 5.

The method in this paper uses $c'_{x,y}$ to replace the gray value of the (x, y) point, so as to achieve the purpose of filtering and retaining the boundary. The definition of $c'_{x,y}$ is shown in

$$c'_{x,y} = \begin{cases} \max(W_{x,y}), & b_{x,y} > T, \\ \text{mean}(W_{x,y}), & \text{others.} \end{cases} \quad (3)$$

Here, T is the boundary judgment threshold, which is selected as 5. If b is greater than the set boundary judgment threshold, the point may be a contour and is replaced by the local maximum; if the intensity of the point is smaller than the threshold, the point is replaced by the mean value of the window area. As shown in Figure 6, the algorithm based on radial region filtering can filter the cervical cell image more effectively and preserve the cell boundary. The overlapping cytoplasmic boundary has a higher resolution after filtering by the regional filtering algorithm.

2.1.2. Gradient Algorithm Based on the Radial of the Nucleus. The contour corresponds to a series of brightness jumps, and the gradient of the image can be used to extract the target contour. There are many boundary gradient algorithms, such as gradient algorithm, Prewitt operator, Sobel operator, etc. [30]. Figure 6 shows the boundary gradient image obtained by the gradient algorithm, Prewitt operator and Sobel operator. As can be seen from Figure 6, the Sobel operator gives a stronger gradient than the other two algorithms. Therefore, this paper uses the Sobel operator based on the radial region of nucleus to generate the gradient map of the image.

This paper proposes a gradient map calculation method based on the radial region Sobel algorithm, which is used to calculate the gradient image f_g of the denoised image. The algorithm utilizes Sobel gradient operators in different directions to enhance the cell boundary. Figure 7 shows the $5 * 5$ Sobel convolution operators in four different directions.

The radial region gradient of point (x, y) is defined as formula (4). In the formula, $*$ represents the convolution operation, and the calculation method of i is shown in formula (4). The maximum gradient and the minimum gradient of the radial region are g_{Max}^i and g_{Min}^i respectively, and the definition of the gradient of each pixel in this region is shown in formula (5). The gradient image of the specified nucleus is f_g , and the definition of f_g is shown in formula (6).

$$g_{i,j}^i = |G(i) \oplus W_{i,j}|, \quad (4)$$

$$f_g^i = 255 \times \frac{(g_{i,j}^i - g_{Min}^i)}{(g_{Max}^i - g_{Min}^i)}, \quad (5)$$

$$f_g = \cup f_g^i, i \in \{1, 2, 3, 4\}. \quad (6)$$

The filtering and gradient calculation method based on the radial region proposed in this paper can remove the

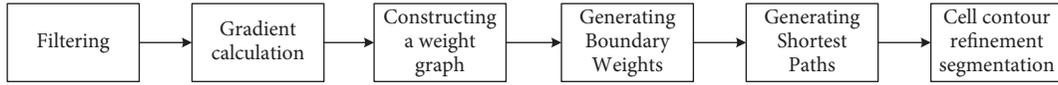


FIGURE 2: The flowchart of the cervical cell segmentation method.

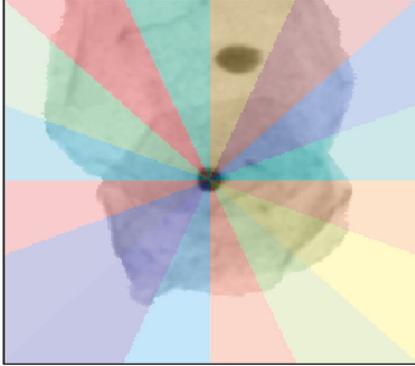


FIGURE 3: Radial division of the cell.

noise in most of the images under the condition of retaining the specified cell boundary. Compared with the methods of literature [2, 33], it has better results, as shown in Figure 8.

2.2. Overlapping Cell Segmentation of Cervical Cytology Images

2.2.1. Constructing a Weight Graph Based on Candidate Boundary Points.

After specifying a cell in the cell clump, the researchers can get a gradient image f_g obtained by the radial region gradient algorithm and then search candidate contour points in the f_g . The weight map of the cell is constructed using the candidate pixel points. For a given cell, there is one and only one true contour point in any radial direction. In general, the contour points have the largest gradient values in the radial direction. However, due to the overlapping of cells and other reasons, the maximum value in the radial direction may not be the real contour point. Therefore, it is necessary to find the pixel points with larger gradient values in the radial direction as candidate contour points, and then find the true cell contour in the candidate contour points.

For the cell C_i in the cell clump, the pixel point on the radial with the nucleus as the centre and θ as the direction is defined as p_k^θ , and the discrete points on the radial can be defined as $l^\theta = p_k^\theta, k = 1, 2, \dots$. In order to obtain the candidate contour points on the radial l^θ , the researchers firstly arrange the pixels on the radial in descending order of intensity value, and take the first 1/3 of the points as the primary selection points of the candidate contour points. In fact, candidate contour points are contained in these pixel points. If a candidate pixel has a large gradient value, the surrounding pixels may also have a large gradient value. The Linkage algorithm [35] is used to cluster the coordinates of candidate contour points. The point with the largest gradient value in each category in the clustering result is selected as the candidate cell contour position. The candidate contour

positions of the gradient image f_g are shown as red marks in Figure 9(a).

It can be seen from the candidate points in the gradient map that each candidate cell contour point in the radial direction contains the real contour point. Next, the researchers construct these candidate contour points into a weight graph, and use the path search algorithm of the weight graph to complete the rough segmentation of the cell boundary.

In order to complete the graph construction of the candidate contour points, the starting point of the candidate contour points is the specified seed contour points. When there is a seed contour point on a radial line segment, other candidate contour points in the radial direction are deleted. The following three steps are used to determine whether a candidate contour point is a seed contour point:

- (1) Detect the candidate contour points located on the boundary of the cell clump
- (2) Connect the point to the centre point of the nucleus, if any point on the connection line is in the background, exclude the point
- (3) Calculate the mean distance of all the seed contour points from the centre of the nucleus, and only retain the seed contour points whose distance from the centre of the nucleus is less than the mean distance of the contour points

The researchers can obtain seed contour points by using the above three steps. The seed contour points are marked in green as shown in Figure 8(b). Next, based on the obtained candidate contour points and seed contour points, a cell coarse segmentation weight graph is constructed. The weight graph construction is mainly divided into four steps:

- (1) Construct N_g column empty weight graph G , each column in the weight graph G corresponds to a radial direction of the cervical cytology image
- (2) Select any seed contour point in the image as the starting contour point, and add a node to the first column of graph G
- (3) Starting from the starting contour point, select the three candidate contour points closest to it in the next radial direction, and add three nodes to the next column of the graph G
- (4) And so on, until the N_g column nodes in the weight graph G are all added

2.2.2. Generating Boundary Weights and Shortest Paths.

The generation of weights between nodes in the weight graph is the core step in solving the optimal contour problem using graph methods. The following constructs the weights between nodes in the weight graph around the attributes of

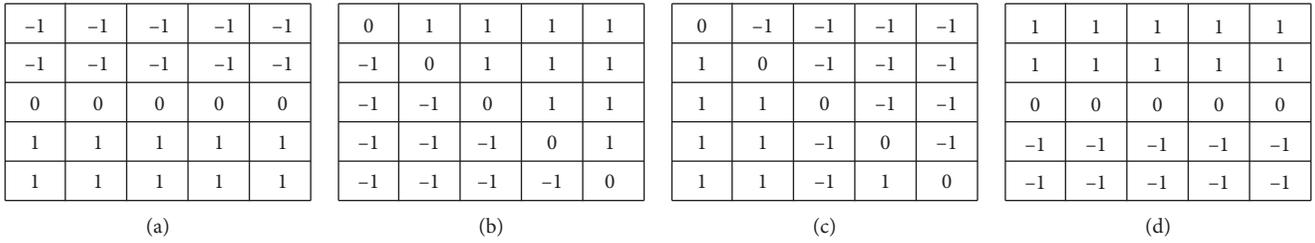


FIGURE 4: Denoising kernel function. (a) Denoising kernel function kernel (1). (b) Denoising kernel function kernel (2). (c) Denoising kernel function kernel (3). (d) Denoising kernel function kernel (4).

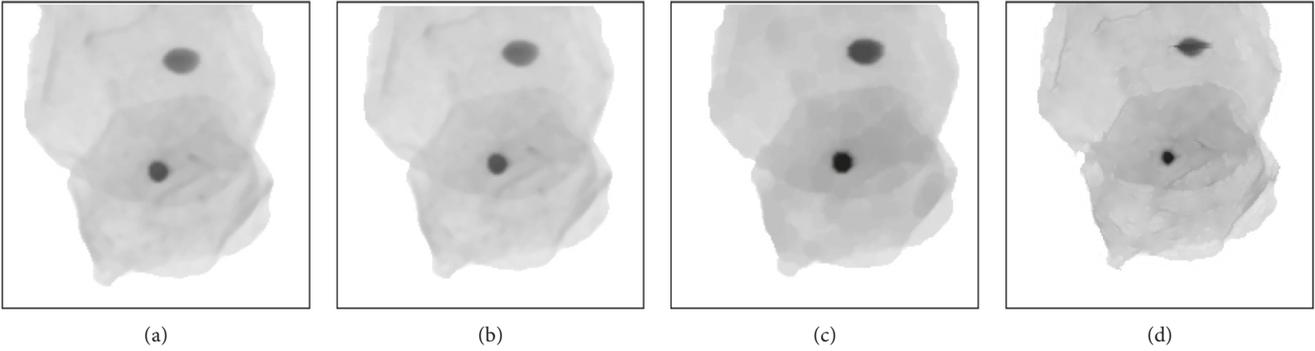


FIGURE 5: Comparison of filtering results of different methods. (a) Original image. (b) Literature [2] results. (c) Literature [33] results. (d) Our results.

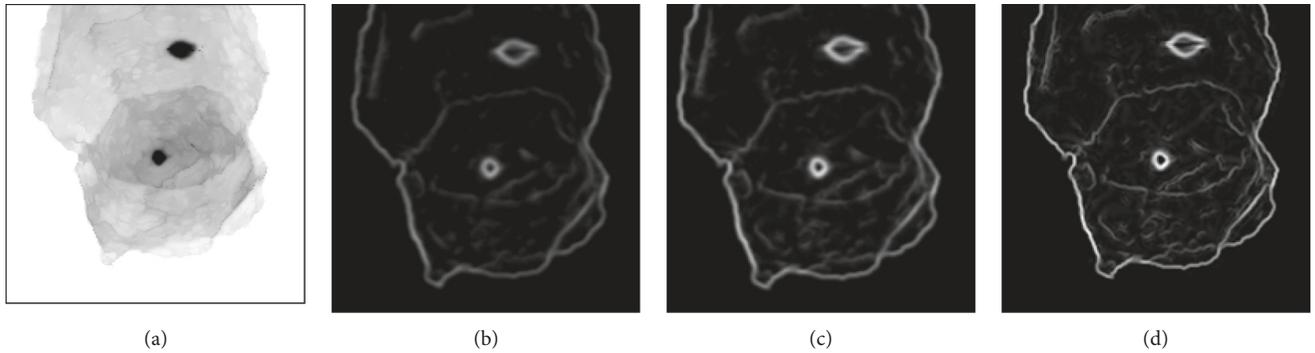


FIGURE 6: Gradient images based on different methods. (a) Original image. (b) Gradient operator. (c) Prewitt operator. (d) Sobel operator.

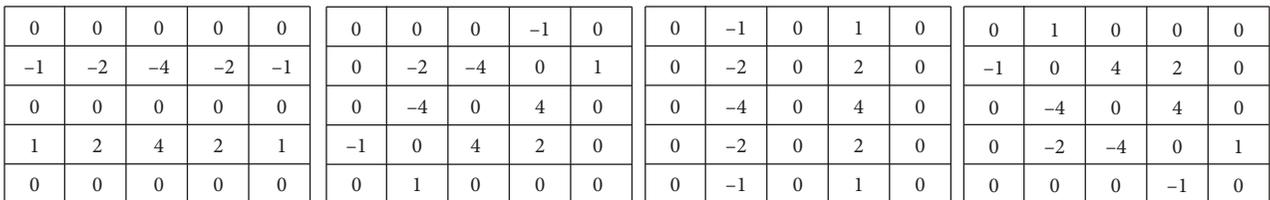


FIGURE 7: Sobel convolution operator in different directions.

contour points and the attributes of line segments between contour points.

Although overlapping cells, uneven illumination, and artefact can degrade image quality, most cell boundaries are still identifiable. By observing the contour points in the

image, it can be seen that the intensity and gradient of different candidate contour points located on the boundary have a certain similarity. The radial gradient of the cell contour points and the window mean are used to define the weights between nodes in the weight graph.

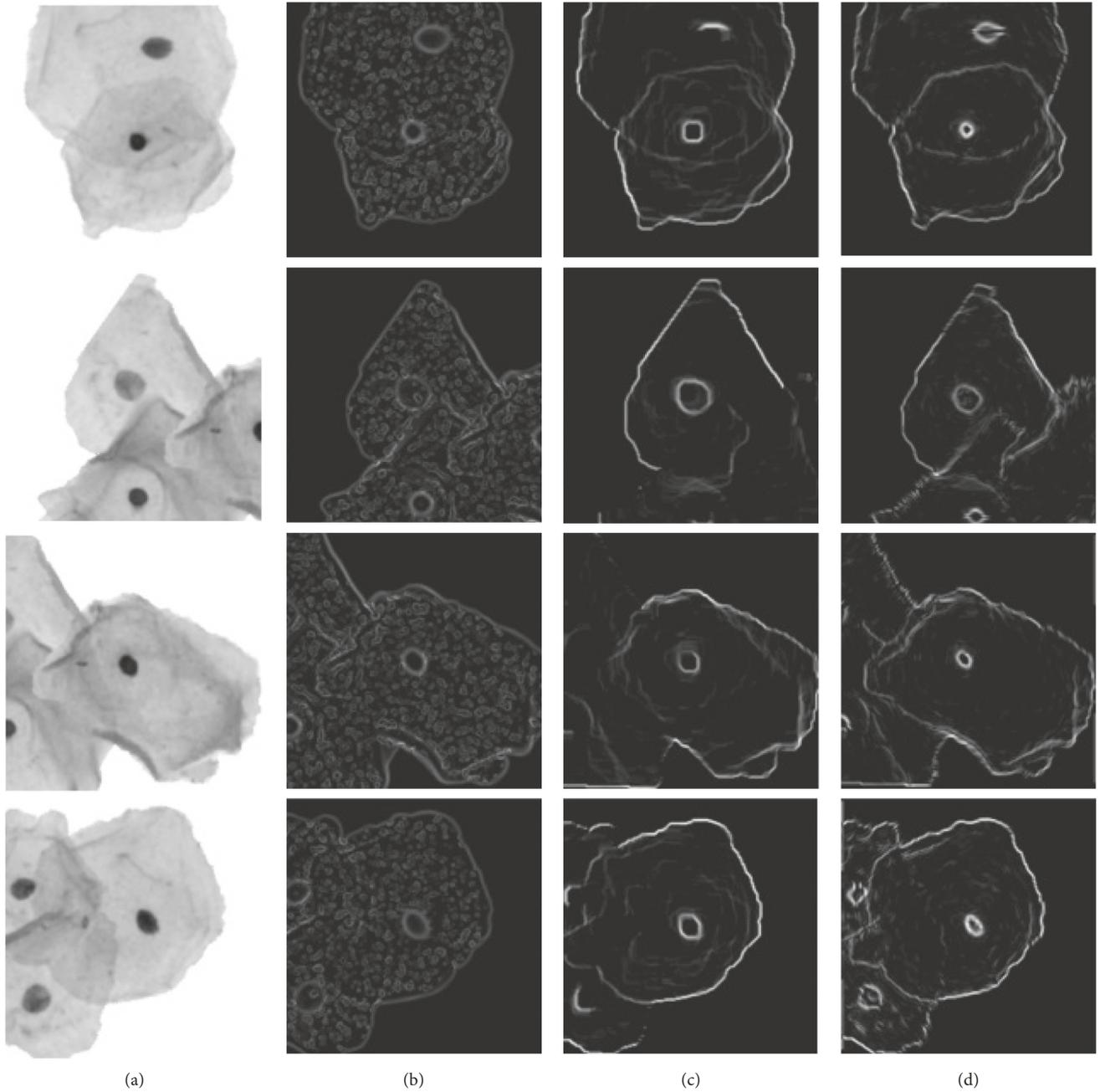


FIGURE 8: Comparison results of overlapping cell denoising and gradient extraction. (a) Original image. (b) Results of [2]. (c) Results of [33]. (d) Our results.

Define $I(x, y)$ as the pixel value of the pixel at the (x, y) position, (x_c, y_c) as the centre of the specified nucleus; in the radial direction θ , the candidate contour point (x_b^θ, y_b^θ) is the farthest point from the center of the nucleus; l^θ is the line segment between the centre of the nucleus (x_c, y_c) and the point (x_b^θ, y_b^θ) . Figure 10 shows an example of two l^θ in different directions.

In this paper, the Bresenham algorithm [36] is used to obtain the coordinate points $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$ on the (x_c, y_c) and (x_b^θ, y_b^θ) line segments and the radial direction difference $rd(x_i, y_i)$ of the coordinate point (x_i, y_i) along l^θ is defined as shown as follows:

$$rd(x_i, y_i) = I(x_{i-1}, y_{i-1}) - I(x_i, y_i). \quad (7)$$

The radial gradient of the coordinate point (x_i, y_i) is shown as follows:

$$rg(x_i, y_i) = \frac{|rd(x_i, y_i)| + |rd(x_{i+1}, y_{i+1})|}{2}. \quad (8)$$

Figure 11(a) shows the gradient intensity distribution in the radial direction. In order to better measure the intensity similarity between contour points, the gray window mean is used to replace the intensity value of the contour points, and the definition of the window mean M is shown as follows:

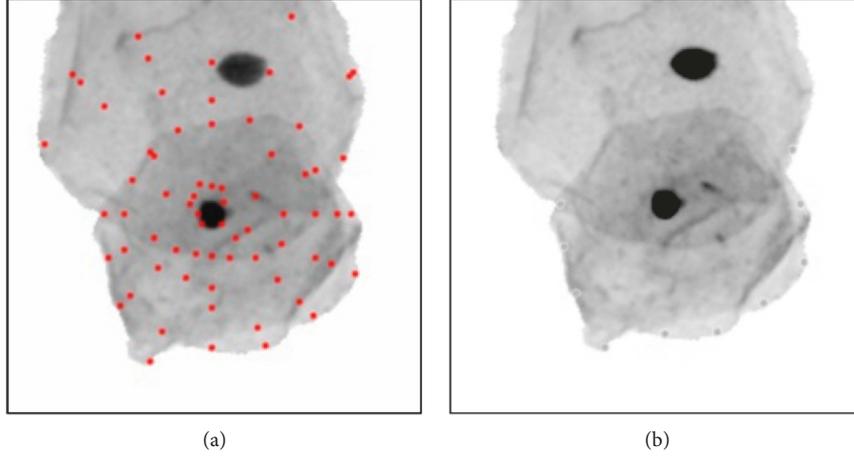


FIGURE 9: Candidate contour positions and seed contour points. (a) Candidate contour positions. (b) Seed contour points.

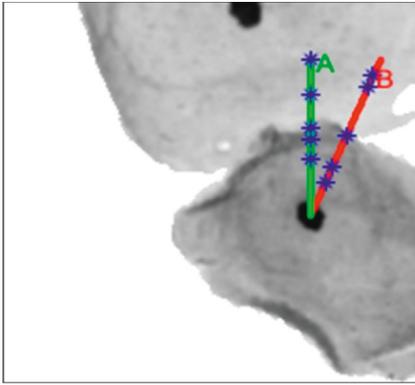


FIGURE 10: Radial segments in different directions.

$$M(x_i, y_i) = \frac{1}{N^2} \sum_{(x,y) \in W} I(x, y). \quad (9)$$

In the formula, W is the $N \times N$ neighbourhood window centred at the (x_i, y_i) point, and $I(x, y)$ is the intensity value of the pixel in the window. Figure 11(b) shows the radial window mean intensity distribution.

For two connected nodes in the weight graph, the smaller the difference between their radial gradient values and the mean, the more likely they belong to the same cell. The gradient-window mean weight w_b is defined as shown in formula (10). The intensity distribution of the gradient-window mean is shown in Figure 11(c).

$$w_b = \left| rg(x_i^\theta, y_i^\theta) - rg(x_j^{\theta+1}, y_j^{\theta+1}) \right| \times \left| M(x_i^\theta, y_i^\theta) - M(x_j^{\theta+1}, y_j^{\theta+1}) \right|. \quad (10)$$

The distance between candidate contour points is a very important condition to measure the similarity between contour points. In general, the smaller the distance between two contour points in adjacent radial directions, the greater the possibility that these two points belong to the same cell. As shown in Figure 12, it is obvious that for node C, it is closer to the contour point of C than B. In fact, both the ‘‘A contour point’’ and the ‘‘C contour point’’ lie on the contour

of the specified cell. There are two main advantages to using distance weights: (1) distance weights reduce the possibility of contours moving away from their corresponding nuclei; (2) distance weights try to keep the cytoplasmic boundary as smooth as possible. In this paper, the distance attribute between candidate contour points is selected as the distance weight in the weight graph. The distance weight between nodes is defined as

$$w_d = \sqrt{(x_i^\theta - x_j^{\theta+1})^2 + (y_i^\theta - y_j^{\theta+1})^2}. \quad (11)$$

The contour point weight w_p between adjacent nodes in the weight graph can be obtained from the gradient-window mean attribute and distance attribute. The definition of w_p is shown as follows:

$$w_p = \alpha w_b + \beta w_d. \quad (12)$$

Here, α and β are the weight factors of the gradient mean attribute item and the distance attribute item, respectively. Before w_p is combined, both w_b and w_d are normalized to $[0, 1]$.

The similarity between contour points has certain limitations, which can only reflect the local features of contour points, and cannot extract sufficient cytoplasmic contour features. Therefore, the attribute of line segment between contour points is further adopted as a supplement to the attribute of contour point.

The contour point line segment refers to the line segment between the candidate contour point (x_i^θ, y_i^θ) on the line segment l^θ in the radial direction of θ and the candidate contour point $(x_j^{\theta+1}, y_j^{\theta+1})$ on the line segment $l^{\theta+1}$. For most cell boundaries, the closer it is to 90 degrees from the radial direction, the more likely it is to be a true cell outline. Therefore, the angle between the radial direction and the line segment is defined as the direction attribute, and the definition is shown as follows:

$$w_{dir} = 1 - \frac{(x_i^\theta - x_c, y_i^\theta - y_c) \cdot (x_j^{\theta+1} - x_i^\theta, y_j^{\theta+1} - y_i^\theta)}{\sqrt{(x_i^\theta - x_c)^2 + (y_i^\theta - y_c)^2} \times \sqrt{(x_j^{\theta+1} - x_i^\theta)^2 + (y_j^{\theta+1} - y_i^\theta)^2}} \quad (13)$$

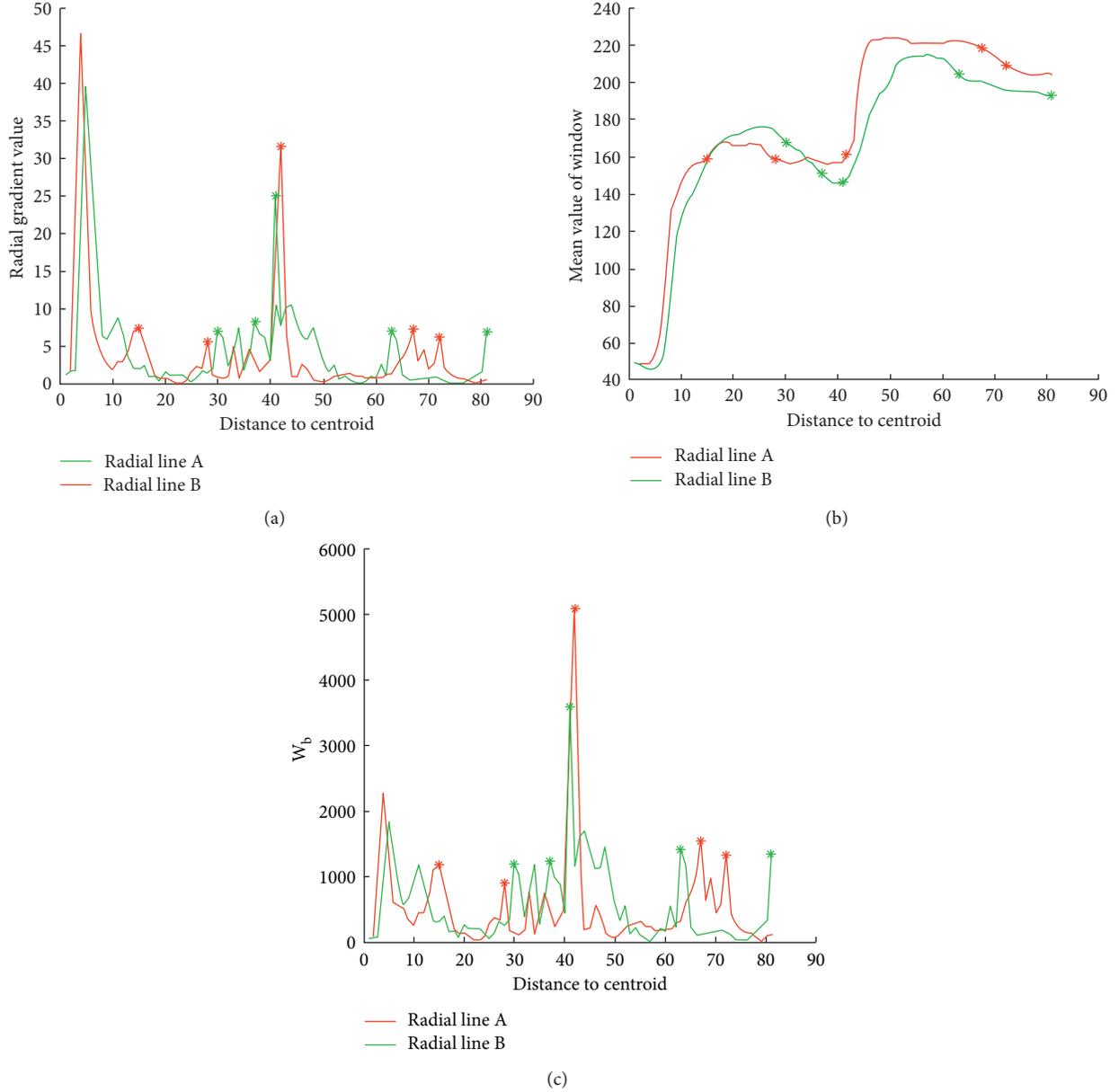


FIGURE 11: Intensity distribution of different attributes in the radial direction. (a) Radial gradient intensity distribution. (b) Radial window mean intensity distribution. (c) Radial gradient-window mean intensity distribution.

By observing the cell image, it can be seen that the cell contour has a significant gradient, and in the gray domain of the image, the boundary region has a large variance. Therefore, the gradient and variance are used to define the attribute weights of the contour segment regions. The contour segment area R_b is obtained by performing morphological operations on the contour point line segment, as shown as follows:

$$R_b = l_b \oplus SE. \quad (14)$$

Among them, SE represents a structural element with a radius of 3 and \oplus is a morphological expansion operation. The definition of the contour segment region attribute w_{R_b} is shown as follows:

$$w_{R_b} = \sigma \left(\frac{1}{N} \sum_{i=1}^N g_i \right)^2. \quad (15)$$

Here, g_i represents the gradient value of the i pixel; N represents the number of pixels in R_b ; σ represents the standard deviation of the area.

Through the analysis and definition of the above-mentioned contour segment direction and area attributes, this paper defines the contour segment weight w_R as shown as follows:

$$w_R = w_{di} r e^{-w_{R_b}}. \quad (16)$$

Finally, combine the attribute weight of the candidate contour point and the contour segment to obtain the weight

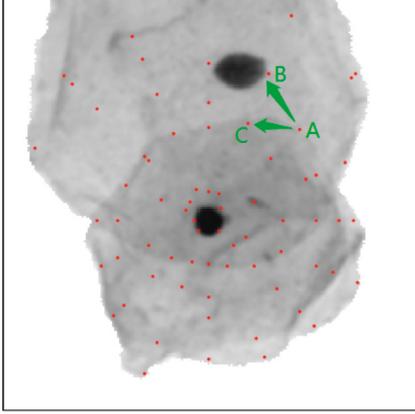


FIGURE 12: Schematic diagram of distance attribute weight.

w_c between the connected nodes in the weight graph. The definition of w_c is shown as follows:

$$w_c = w_p w_R. \quad (17)$$

After the weight graph is constructed, the Dijkstra dynamic programming algorithm is used to find the shortest path in the graph. The shortest path in the graph corresponds to the rough segmentation contour in the cell image, as shown in Figure 13(a).

2.2.3. Cell Contour Refinement Segmentation. Since the rough segmentation boundary is only the connection between the candidate contour points in the radial direction, it can only roughly describe the boundary of the specified cell. In order to obtain accurate cytoplasmic contours, this paper uses the distance regularized level set evolution (DRLSE) model [31] to fine-tune the obtained rough segmentation boundaries of cells. The rough cytoplasmic segmentation contour obtained is used as the initial value of the DRLSE level set algorithm.

The energy functional of the DRLSE level set algorithm can be approximately expressed as $\varepsilon(\phi)$

$$\begin{aligned} \varepsilon(\phi) = & \mu \int_{\Omega} P(|\nabla\phi|) dx + \lambda \int_{\Omega} g \delta(\phi) |\nabla\phi| dx \\ & + \alpha(\nu) \int_{\Omega} g H(-\phi) dx. \end{aligned} \quad (18)$$

The minimum value of the energy function is obtained by solving the gradient flow, as shown as follows:

$$\begin{aligned} \frac{\partial\phi}{\partial t} = & \mu \operatorname{div} \left(d_p(|\nabla\phi|) \nabla\phi \right) + \lambda \delta_{\varepsilon}(\phi) \operatorname{div} \left(g \frac{\nabla\phi}{|\nabla\phi|} \right) \\ & + \alpha(\nu) g \delta_{\varepsilon}(\phi). \end{aligned} \quad (19)$$

When the energy functional reaches a minimum, the level set iteration stops. Figure 13(b) shows the refined cytoplasmic segmentation result of the DRLSE algorithm. It can be seen that the coarsely segmented cell boundary has evolved to the vicinity of the real cytoplasmic boundary, and

the red boundary is the refined cytoplasmic segmentation boundary obtained by the DRLSE algorithm.

3. Results

3.1. Data. In order to evaluate the performance of the proposed overlapping cervical cytology segmentation algorithm, the proposed algorithm was evaluated using the datasets published in the first and second ‘‘ISBI Overlapping Cervical Cell Segmentation Challenge.’’ The samples in the dataset contained different numbers of cervical cells, and the cells varied in contrast, texture, and overlap between adjacent cells. Table 1 summarizes the datasets published by the two challenges.

The first ISBI Cervical Overlapping Cell Challenge was published in two separate datasets containing 945 synthetic images with a resolution of $512 * 512$. For the first time, the challenge organizing committee announced 45 synthetic images and 90 test images. Among them, 45 synthetic images are used to adjust and train the parameters of the segmentation algorithm and 90 test images are used to evaluate the algorithm performance. The second published dataset contains 810 synthetic cervical overlay images that were used to evaluate the performance of the algorithms submitted by the contestants.

The second ISBI Cervical Overlapping Cell Challenge dataset contains 17 sets of cervical cell images from different cervical cell samples, of which the training set is 8 and the others are the test set. Each set of cervical cell images contains 20 stacked images and 1 EDF (Extended Depth of Field) image [37], and the resolution of each image is $1024 * 1024$. Each set of images contains 20–60 cells distributed in different cell clumps. The number of cells in the clump and the overlap ratio between adjacent cells are different. For each set of stacked images, each image corresponds to a different focal plane of the same sample. The depth of the different focal planes under the microscope is approximately 1 micron. The test images marked the cytoplasmic boundaries by manual annotation. The overlapping cell segmentation algorithm proposed in this paper is evaluated on test images in different datasets.

3.2. Cytoplasm Segmentation Evaluation Method. The cervical cytology segmentation evaluation methods specified by the ISBI challenge include: pixel-based dice similarity coefficient ($DS C$), pixel-based true positive rate (TP_p), pixel-based false positive rate (FP_p), and object-based false negative rate (FN_o). If the $DS C$ metric value of the overlapping area between the obtained cell segmentation region O_{seg} and the manually labelled (gold standard) cell region O_{GT} is higher than the specified threshold, the segmentation algorithm is success to identify the cell. DSC is shown as follows:

$$DSC = \frac{2|O_{GT} \cap O_{seg}|}{|O_{GT}| + |O_{seg}|}. \quad (20)$$

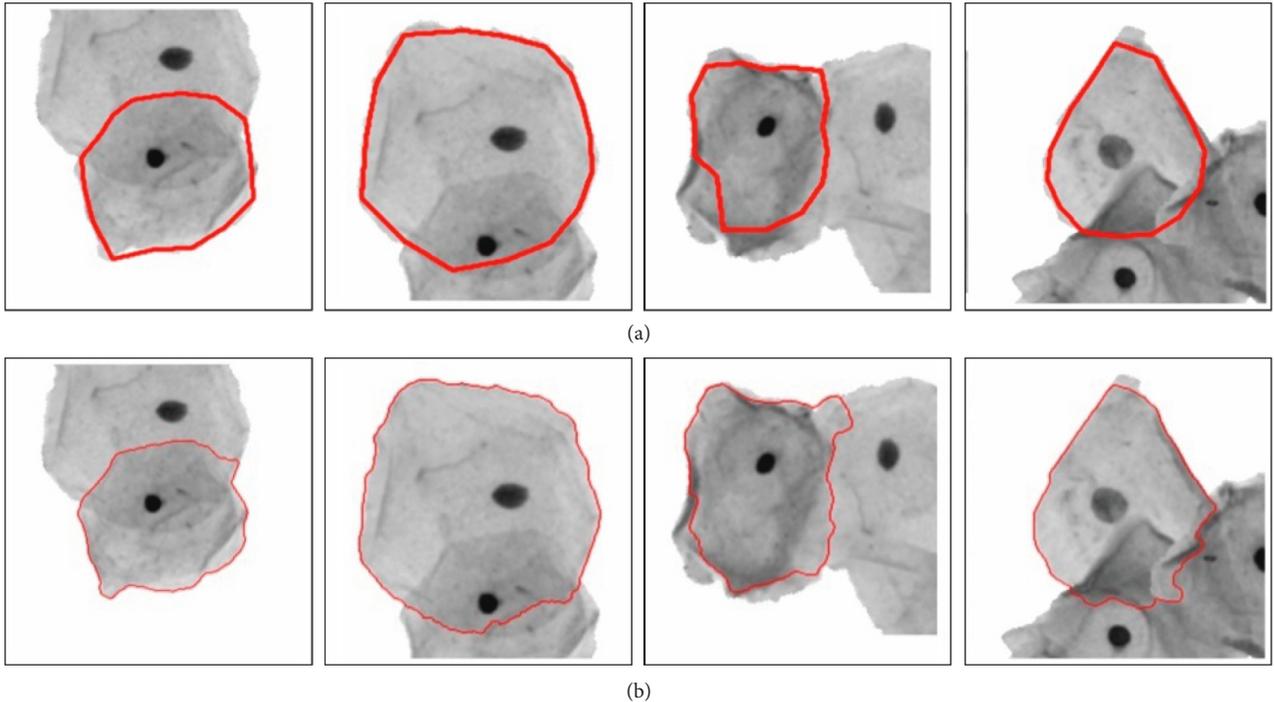


FIGURE 13: Coarse segmentation and fine segmentation results. (a) Coarse segmentation contour. (b) Refined segmentation contour.

TABLE 1: ISBI overlapping cervical cell segmentation challenge dataset.

Challenge	Dataset	Type	Resolution	Number
ISBI2014	Training set	Synthetic images	512 * 512	45
	Test set A			90
	Test set B			810
ISBI2015	Training set	EDF images	1024 * 1024	8
	Test set			9

Here, $|\cdot|$ represents the number of pixels in the region. The ISBI Challenge specified threshold $DS C = 0.7$ for correct cytoplasmic identification. For the cells in the manually labelled (gold standard) dataset, if there is no $DS C$ value between the segmented cells and the manually labelled cells in the segmentation result greater than 0.7, the segmentation algorithm is said to have not segmented the cell correctly.

The cytoplasmic evaluation standard FP_p provided by ISBI Challenge is false positive rate based on pixel points. Due to the small cell size relative to the image size, previous studies [4, 15, 23, 28, 37] have achieved segmentation results with FP_p less than 0.01 for different datasets. It can be seen that the FP_p metric cannot effectively measure the performance of different cytoplasmic segmentation algorithms. Therefore, this paper adopts the object-based false recognition rate FD_o proposed in [24] to replace the FP_p used in the ISBI challenge. FD_o represents the proportion of misidentified cytoplasm among the objects identified by the segmentation algorithm. The combination of FD_o and FN_o can effectively estimate the cytoplasm identification performance of different algorithms. This paper uses the geometric median metric GM of $DS C$ and FN_o to measure the

overall performance of the segmentation algorithm in terms of accuracy and recognition rate. The geometric median GM [24] is defined as

$$GM = \sqrt{DSC(1 - FNR_o)}. \quad (21)$$

3.3. Cytoplasm Segmentation Results. In this paper, the parameters of the DLRSE algorithm for cell nucleus refinement segmentation are set as follows: the value of internal iterations $iter_inner = 10$; the value of outer $iter_outer = 5$; the weight of the regular term $\mu = 0.04$; the time step $\Delta t = 5$. The weight parameters of the contour length functional $L_g(\phi)$ and area energy functional $A_g(\phi)$ of the zero level set are $\alpha = -4$, $\lambda = 6$, respectively.

For the weight value α and β of the distance attribute, the researchers need to find the best combination of α and β such that the overlapping cell segmentation algorithm achieves the highest GM metric value on the training set. Within the range of 5 sets of parameters α (0.2, 0.5, 0.8, 1, 1.5) and 10 sets of parameters β (0.2, 0.4, 0.6, ..., 2), the optimal combination (0.8, 1.4) of α and β was found through experiments to obtain the maximum GM . The experimental results of the highest GM , $DS C$ and the lowest FN_o and FD_o are shown in Table 2.

The test dataset published by the ISBI2014 challenge contains 810 images of cells with different cell numbers and overlapping ratios. The 810 images in the test set were divided into 45 subsets according to the cell number and overlap ratio, and each subset contained 18 synthetic images with the same cell number and overlap ratio. Based on the 45 subsets of the ISBI2014 test set, the performance of

TABLE 2: Segmentation results with different combinations of α and β .

α	β	<i>DSC</i>	TP_p	FN_o	FD_o	<i>GM</i>
0.2	1.4	0.91	0.93	0.061	0.07	0.924
0.5	1	0.9	0.936	0.045	0.051	0.927
0.8	1	0.881	0.935	0.047	0.042	0.916
0.8	1.6	0.882	0.935	0.046	0.040	0.917
0.8	1.4	0.907	0.934	0.044	0.05	0.931

The bold font is the best result of the metric in the comparison results.

segmentation algorithm with different cell numbers and overlapping rates can be evaluated. The cell image segmentation results of our algorithm with $DS C=0.7$ for different cell numbers and overlapping rates are shown in Table 3. It can be clearly seen from it that when the overlap rate between cells is between $[0, 1]$, our algorithm can identify all cells and obtain $DS C$ not less than 0.95. When the overlap rate is in the range of $[0.1, 0.2]$, the algorithm can also identify all cells when the number of cells is less than 6; when the number of cells is greater than 6, only individual cells are not identified. Therefore, when the overlap ratio is less than 0.2, our algorithm can ignore the number of overlapping cells in the image, and the cell segmentation metric results TP_p and $DS C$ are higher than 0.95 and 0.93, respectively. The above analysis shows that the overlapping cell algorithm in this paper can not only accurately identify the cells in the overlapping cell images with the overlapping rate in the range of $[0, 0.2]$ but also accurately segment them.

When the overlap ratio is in $[0.2, 0.4]$, the proportion of misidentified cell numbers increases slightly with increasing cell numbers. When the number of overlapping cells is less than 8, the segmentation accuracy indicators $DS C$ metrics and TP_p metrics are both higher than 0.86. When the overlap rate is in $[0.4, 0.5]$, the correct identification rate of overlapping cells decreases significantly with the increase of the number of cells. This is because when the cells overlap, the resolution of the cell boundary is significantly reduced, which makes the algorithm cannot accurately identify cells.

It can be seen from Table 3 that when there are only 2 or 3 cells in the image, the FN_o metric value of our algorithm is less than 0.07. When the overlap rate in the image is in the range $[0, 0.3]$, the algorithm in this paper can identify all overlapping cells. Although some cells with overlapping rates in the range of $[0.4, 0.5]$ were not correctly identified, the object-based false positive rate of our algorithm is close to 0.1. This shows that when the number of cells is small, most overlapping cells can be correctly identified even if the overlapping rate is high.

When the overlap ratio between cells was in the range $[0.4, 0.5]$, both $DS C$ and TP_p metrics for correctly identified cells were higher than 0.8. It can be seen that even in the overlapping cell clumps with a high overlap rate and a large number of cells, our algorithm can still perform well in the identification and segmentation of cells in the overlapping cell clump. Figure 14 can more clearly show the change trend of the segmentation results with the number of cells and the overlap rate.

For the ISBI2014 and ISBI2015 challenge datasets, some scholars have proposed different overlapping cell segmentation methods. Next, based on the test dataset of the ISBI challenge, the proposed method and other overlapping cervical cell segmentation methods are compared quantitatively and qualitatively.

First, compare the algorithm in this paper with the algorithms mentioned in other literature based on the test set of ISBI2014 challenge, including the ISBI2014 challenge champion and runner-up algorithms: Ushizima et al. [16], Nosrati and Hamarneh [38], and some algorithms after the challenge [14, 23, 28, 37]. Then, based on the 9 EDF images of the ISBI2015 test set, our algorithm is quantitatively and qualitatively compared with the ISBI challenge winner and runner-up algorithms: Phoulady et al. [23] and Ramalho et al. [39] and some recent algorithms [22, 28, 37, 40].

As shown in Table 4, based on the test set of ISBI2014 challenge, the bold font is the best result of the metric in the comparison results. The test set contains 810 cervical cell images with different numbers and overlapping ratios. From the metrics in Table 4, it can be seen that the previous studies on cytoplasmic segmentation algorithms have obtained high metric values, and Phoulady's algorithm [24] with the best performance is 0.901.

For this dataset, the TP_p metric values of the previous algorithms are all greater than 0.8. The lowest TP_p value of Lu's algorithm is 0.805, and the highest is Tareef's three-step watershed algorithm, which has a TP_p value of 0.94. Combining the DSC and TP_p metrics, it can be seen that earlier algorithms have higher segmentation accuracy for correctly identified cytoplasm. DSC and TP_p metric values of the cell segmentation of our algorithm are 0.904 and 0.927, respectively, and the DSC metric value is higher than the highest Phoulady's algorithm [24] by 0.33%. However, for the TP_p metric, our algorithm is slightly lower than Tareef's three-step watershed algorithm but higher than the other six algorithms. It can be seen from the comparison results of $DS C$ and TP_p metric values that the segmentation method in this paper has a greater advantage in segmentation accuracy than other cervical cell segmentation algorithms.

FN_o and FD_o metric values characterize the cell identification ability of the cell segmentation algorithm. From the FN_o values of different algorithms provided in Table 4, it can be seen that the metric values of different algorithms are quite different. Under the FN_o metric, the proposed algorithm and Phoulady's algorithm [24] have great advantages over other methods, and the proposed algorithm obtains the lowest FN_o metric value. Among the algorithms of Ushizima, Lu, and Phoulady that provide FD_o metric values, Phoulady's algorithm [24] achieves the lowest FD_o metric value, and the algorithm in this paper is slightly higher than Phoulady's algorithm. The FD_o metric of our algorithm and Phoulady's algorithm is much lower than the other three methods that provide FD_o metric. Combining FN_o and FD_o measurement results, it can be seen that compared with other methods, the algorithm in this paper and Phoulady [24] have great advantages in overlapping cell identification.

TABLE 3: Segmentation results of different cell numbers and overlap rates (DSC , TP_p , FN_0 , and GM).

	[0, 0.1]	(0.1, 0.2]	(0.2, 0.3]	(0.3, 0.4]	(0.4, 0.5]
2	0.99/0.98/0/0.99	0.97/0.97/0/0.98	0.94/0.95/0/0.96	0.9/0.92/0.01/0.94	0.88/0.9/0.058/0.9
3	0.99/0.99/0/0.99	0.97/0.97/0/0.98	0.94/0.95/0/0.96	0.9/0.92/0.01/0.94	0.85/0.9/0.07/0.89
4	0.96/0.97/0/0.98	0.99/0.98/0.01/0.99	0.92/0.97/0.04/0.96	0.88/0.9/0.01/0.93	0.87/0.88/0.06/0.9
5	0.97/0.97/0/0.98	0.96/0.96/0/0.98	0.93/0.96/0/0.96	0.89/0.9/0.02/0.93	0.86/0.89/0.08/0.89
6	0.96/0.96/0/0.98	0.95/0.95/0/0.97	0.92/0.95/0.01/0.95	0.86/0.88/0.04/0.91	0.86/0.86/0.09/0.88
7	0.95/0.97/0/0.97	0.95/0.95/0.02/0.96	0.92/0.95/0.04/0.94	0.87/0.9/0.02/0.92	0.86/0.86/0.04/0.9
8	0.97/0.97/0/0.98	0.96/0.96/0.01/0.97	0.91/0.95/0.07/0.92	0.83/0.87/0.1/0.86	0.82/0.85/0.12/0.85
9	0.96/0.97/0/0.98	0.94/0.96/0.02/0.96	0.9/0.95/0.03/0.93	0.84/0.87/0.07/0.88	0.81/0.85/0.16/0.82
10	0.96/0.98/0/0.98	0.93/0.96/0.01/0.96	0.88/0.95/0.06/0.91	0.83/0.88/0.14/0.84	0.81/0.86/0.16/0.82

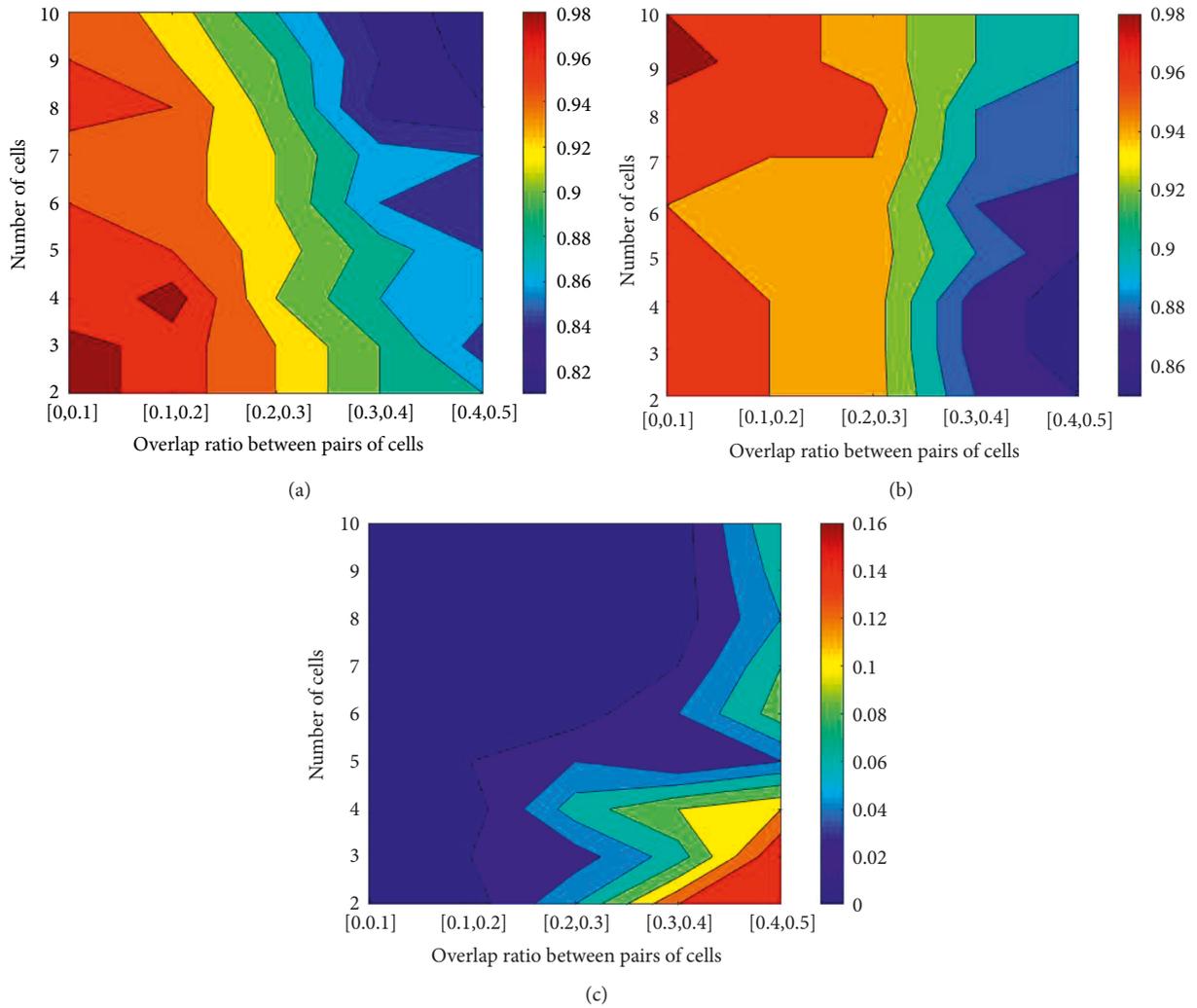


FIGURE 14: Visualization results of Table 3. (a) Visualization results of DSC . (b) Visualization results of TP_p . (c) Visualization results of TP_o .

TABLE 4: Different method results based on the ISBI2014.

	DSC	TP_p	FN_0	FD_0	GM
Ushizima	0.872	0.841	0.267	0.152	0.799
Nosrati	0.871	0.875	0.11	—	0.880
Lu	0.893	0.805	0.315	0.146	0.782
Lee	0.897	0.882	0.137	—	0.880
Phoulady [23]	0.883	0.902	0.087	0.142	0.898
Phoulady [24]	0.901	0.906	0.07	0.041	0.915
Tareef	0.898	0.946	0.161	—	0.868
Ours	0.904	0.927	0.068	0.050	0.918

TABLE 5: Different method results based on the ISBI2015.

Methods	DSC	TP_p	FN_o	GM
Phoulady	0.831	0.927	0.408	0.701
Ramalho	0.856	0.899	0.501	0.654
Lu	0.856	0.892	0.392	0.721
Lee	0.879	0.877	0.434	0.705
Kumar	0.852	0.885	0.361	0.738
Phoulady	0.861	0.874	0.352	0.747
Tareef	0.851	0.94	0.336	0.751
Ours	0.882	0.851	0.318	0.774

The bold font is the best result of the metric in the comparison results.

In order to more comprehensively evaluate the algorithm's performance for overlapping cell detection and segmentation, this paper evaluates the above methods through the metric GM . It can be seen from Table 4 that the GM measurement values of Lu and Ushizima methods are low, 0.782 and 0.799, respectively; the GM measurement values of Nosrati, Phoulady [23], Lee, and Tareef methods are medium, between 0.868 and 0.898. This paper and Phoulady's [24] algorithm have higher GM measurement values of 0.918 and 0.915, respectively.

4. Discussion

In summary, for the ISBI2014 synthetic dataset, under the $DS C$, FN_o , GM metrics, the algorithm in this paper has obtained the best results compared with other algorithms. Under the TP_p and FD_o metrics, our algorithm is only slightly worse than the best performing method. Therefore, for the ISBI2014 dataset with different cell numbers and overlapping rates, the algorithm in this paper has achieved better segmentation results.

Table 5 shows the comparison results of overlapping cell segmentation between our and other algorithms based on the ISBI2015 challenge test set, where the bold font is the best result of this metric. The test set contains 9 EDF images, each image contains multiple cell clumps, and the cell clumps contain different cell numbers and overlapping rates.

On the ISBI2015 dataset, our algorithm achieves the best metric values relative to other methods on $DS C$, TP_p , and GM metrics. For the $DS C$ metric, the previous algorithms have obtained higher metric values, the lowest is Phoulady's [22] algorithm, with a $DS C$ value of 0.831; the highest is Lee's algorithm, with a $DS C$ value of 0.879. The algorithm in this paper is 5.57% and 0.11% higher than the two algorithms under the $DS C$ metric, respectively. In terms of recognition rate, there is a big difference between different algorithms. The FN_o metric value of Ramalho algorithm is only 0.501. In recent years, some algorithms have significantly improved the performance of overlapping cell recognition, such as the algorithms of Phoulady and Tareef. Under the FN_o metric, these two algorithms obtained values of 0.352 and 0.336, respectively, which were 42% and 47% lower than that of Ramalho's algorithm, respectively. Compared with Phoulady et al. [23], the FN_o of Tareef [37] algorithm

decreased by 5% and 10%, respectively. It can be seen that the algorithm in this paper has great advantages in overlapping cell detection. The algorithm in this paper also achieves the best results on GM evaluation metric, which are 3% and 3.5% higher than Phoulady's [23] and Tareef's [37] algorithms, respectively.

For the ISBI2014 and ISBI2015 challenge datasets, our algorithm achieves the best results on the $DS C$, FN_o , and GM metrics. Since the ISBI2015 dataset is a real cervical cytology image, the image quality is more seriously affected by uneven illumination and cell overlap. Therefore, the recognition rate and segmentation accuracy based on the ISBI2015 dataset are lower than the ISBI2014 dataset.

Figure 15 shows the overlapping cytology segmentation results of this paper, manual annotation, and four other algorithms for ISBI2014 synthetic images and ISBI2015 EDF images. As shown in Figure 15(b), Ushizima's overlapping cell segmentation method [16] uses straight line segments to separate adjacent overlapping cells. It is clear that this method cannot obtain accurate boundaries between overlapping cells. This method uses the segmentation results obtained by the watershed algorithm, which cannot describe the true boundaries between overlapping cells. In contrast, Tareef's algorithm [37] solves this problem using an ellipse model and an iterative watershed algorithm. It can be seen from Figure 15(e) that the overlapping boundary between cells can be estimated using the watershed algorithm of the ellipse model.

As can be seen from Figure 15(c), the cells segmented by Nosrati's algorithm [38] are usually located inside the real cells, which cause a high false negative rate and only describe an approximate cell boundary. The method proposed by Lu et al. [14] performs well for the segmentation of synthetic overlapping cell images. However, for EDF images, although the segmentation method can roughly segment overlapping cells, the segmentation results are not ideal, and the oversegmentation and undersegmentation are serious. Among the four qualitatively compared methods, the MPFW method proposed by Tareef et al. [37] has the best segmentation results, and this method achieves better results whether it is synthetic cell images or EDF images. However, because this method adopts the iterative level set method of ellipse prior, the overlapping boundary is replaced by ellipse, which leads to many cases of oversegmentation.

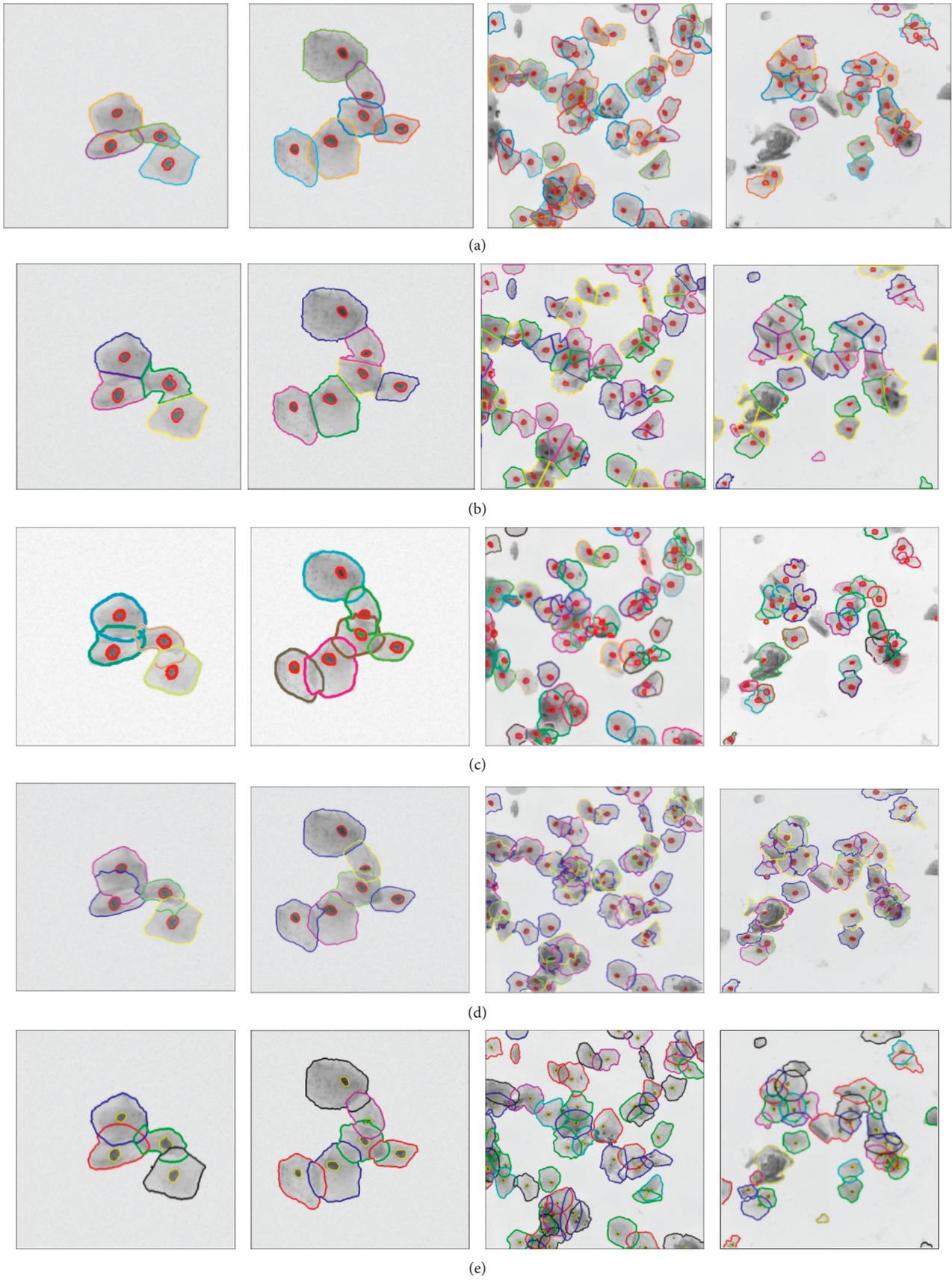


FIGURE 15: Continued.

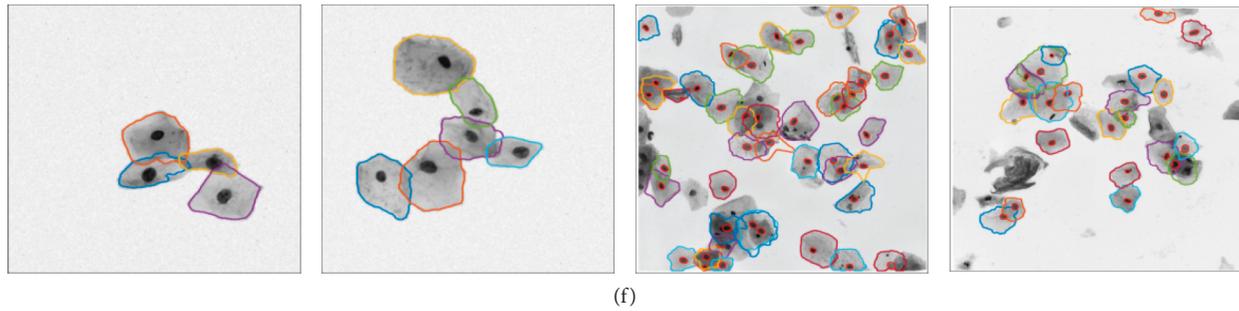


FIGURE 15: Segmentation results of different types of overlapping cervical cell images. (a) Manual labelling (ground truth) results. (b) Segmentation results of Ushizima's algorithm. (c) Segmentation results of Nosrati's algorithm. (d) Segmentation results of Lu's algorithm. (e) Segmentation results of Tareef's algorithm. (f) Segmentation results of our algorithm.

Figure 15(f) is the segmentation result of the algorithm in this paper for synthetic images and EDF images. Observing the segmentation results in this paper, it can be seen that the overlapping cell segmentation algorithm in this paper can successfully detect the cell boundaries located in the overlapping area. By comparing Figure 15(f) with the gold standard in Figure 15(a), it can be seen that the cytoplasmic segmentation results in this paper are very similar to the boundaries in the gold standard. Therefore, the overlapping cervical cell segmentation method proposed in this paper can complete the segmentation of cells in overlapping cell clumps of different types of datasets. The algorithm in this paper has some segmentation errors in EDF images. The reason is that the occlusion between some cells in the EDF image is too severe, and the cells are completely located between overlapping cells, so accurate candidate boundary points cannot be obtained.

5. Conclusions

This paper proposes an overlapping cytology segmentation method for the ISBI2014 synthetic dataset and the ISBI2015 real cervical cell dataset. The noise of overlapping cervical cell images can be significantly suppressed by the proposed cell-based radial region filtering algorithm, and the contrast of cell boundaries in the images is preserved. By generating the weight graph and boundary weights based on the attributes of candidate contour points and contour line segments, a coarse segmentation boundary is obtained. Comparing the algorithm in this paper with the algorithms mentioned in other literature based on the test set of ISBI2014 and ISBI2015 test set, our segmentation method has a greater advantage in segmentation accuracy than other cervical cell segmentation algorithms. Observing the segmentation results in this paper, it can be seen that the overlapping cell segmentation algorithm in this paper can successfully detect the cell boundaries located in the overlapping area. It can be seen from the quantitative and qualitative evaluation results that the overlapping cytoplasm segmentation algorithm in this paper has achieved better segmentation results. Compared with other current overlapping cytoplasm segmentation algorithms, the segmentation results obtained in this paper have great

advantages. Next, the author will study the key technologies of cervical cancer screening based on graph attention neural network.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported in part by the Youth Reserve Talent Support Scheme of Harbin University of Commerce (2020CX22) and Doctoral Research Support Program of Harbin University of Commerce (22BQ29).

References

- [1] M. E. Plissiti and C. Nikou, "On the importance of nucleus features in the classification of cervical cells in Pap smear images[J]," *University of Ioannina*, vol. 7325, pp. 483–490, 2012.
- [2] S. F. Yang-Mao, Y. K. Chan, and Y. P. Chu, "Edge enhancement nucleus and cytoplasm contour detector of cervical smear images[J]," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 2, pp. 353–366, 2008.
- [3] N. M. Harandi, S. Sadri, N. A. Moghaddam, and R. Amirfattahi, "An automated method for segmentation of epithelial cervical cells in images of ThinPrep," *Journal of Medical Systems*, vol. 34, no. 6, pp. 1043–1058, 2010.
- [4] K. Li, Z. Lu, W. Liu, and J. Yin, "Cytoplasm and nucleus segmentation in cervical smear images using Radiating GVF Snake," *Pattern Recognition*, vol. 45, no. 4, pp. 1255–1264, 2012.
- [5] H. S. Hai-Shan Wu, J. Barba, and J. Gil, "A parametric fitting algorithm for segmentation of cell images," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 3, pp. 400–407, 1998.
- [6] T. Chankong, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in Pap smears," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 2, pp. 539–556, 2014.

- [7] M. H. Tsai, Y. K. Chan, Z. Z. Lin, S. F. Yang-Mao, and P. C. Huang, "Nucleus and cytoplasm contour detector of cervical smear image," *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1441–1453, 2008.
- [8] C. Chenyang Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.
- [9] R. Song, F. Wu, G. Wu, L. Liu, D. Zhu, and Z. Zhu, "identification of cancerous cell lines and evaluation of cytotoxic effects on cancer cells with VUV-pims," *International Journal of Advanced Nursing Education and Research*, vol. 2, no. 2, pp. 69–86, 2017.
- [10] A. Bhardwaj, A. Goyal, and A.P Shukla, "Performance comparison of various filters on despeckling of medical ultrasound imaging," *International Journal of IT-based Public Health Management*, vol. 3, no. 2, pp. 9–16, 2016.
- [11] J. Tang, "A multi-direction GVF snake for the segmentation of skin cancer images," *Pattern Recognition*, vol. 42, no. 6, pp. 1172–1179, 2009.
- [12] M. E. Plissiti, C. Nikou, and A. Charchanti, "Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 2, pp. 233–241, 2011.
- [13] Y. Song, L. Zhu, J. Qin, B. Lei, B. Sheng, and K. S. Choi, "Segmentation of overlapping cytoplasm in cervical smear images via adaptive shape priors extracted from contour fragments," *IEEE Transactions on Medical Imaging*, vol. 38, no. 12, pp. 2849–2862, 2019.
- [14] Z. Lu, G. Carneiro, and A. P. Bradley, "An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells[J]," *IEEE Transactions on Image Processing*, vol. 24, no. 4, pp. 1261–1272, 2015.
- [15] 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.
- [16] D. M. Ushizima, A. G. C. Bianchi, and C. M. Carneiro, *Segmentation of Subcellular Compartments Combining Superpixel Representation with Voronoi diagrams[R]*, Lawrence Berkeley National Lab, Berkeley, CA, USA, 2015.
- [17] R. Nock and F. Nielsen, "Statistical region merging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1452–1458, 2004.
- [18] G. W. Zack, W. E. Rogers, and S. A. Latt, "Automatic measurement of sister chromatid exchange frequency," *Journal of Histochemistry and Cytochemistry*, vol. 25, no. 7, pp. 741–753, 1977.
- [19] Y. Yin, Z. Cao, Y. Xu, H. Gao, R. Li, and Z. Mai, "QoS prediction for service recommendation with features learning in mobile edge computing environment," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 4, pp. 1136–1145, 2020.
- [20] Z. Liu, P. Qian, X. Wang, Y. Zhuang, L. Qiu, and X. Wang, "Combining graph neural networks with expert knowledge for smart contract vulnerability detection," *IEEE Transactions on Knowledge and Data Engineering*, 2021.
- [21] T. Wang, "Segmentation of cervical cell cluster by multiscale graph cut algorithm[C]," in *Proceedings of the International Conference on Business Intelligence and Information Technology*, pp. 131–140, Springer, Harbin, China, December 2021.
- [22] H. A. Phoulady, D. B. Goldgof, and L. O. Hall, "An approach for overlapping cell segmentation in multi-layer cervical cell volumes[J]," *The second overlapping cervical cytology image segmentation challenge-IEEE ISBI*, 2015.
- [23] H. A. Phoulady, D. B. Goldgof, and L. O. Hall, "A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images[C]," in *Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 201–204, IEEE, Prague, Czech, April 2016.
- [24] H. A. Phoulady, D. Goldgof, L. O. Hall, and P. R. Mouton, "A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images," *Computerized Medical Imaging and Graphics*, vol. 59, pp. 38–49, 2017.
- [25] M. S. Nosrati and G. Hamarneh, "Segmentation of overlapping cervical cells: a variational method with star-shape prior[C]," in *Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 186–189, IEEE, Brooklyn, NY, USA, April 2015.
- [26] P. Bamford and B. Lovell, "A water immersion algorithm for cytological image segmentation[C]," in *Proceedings of the APRS Image Segmentation Workshop*, pp. 75–79, Sydney, Australia, December 1996.
- [27] A. Tareef, Y. Song, H. Huang et al., "Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling," *Neurocomputing*, vol. 248, pp. 28–40, 2017.
- [28] J. Huang, T. Wang, D. Zheng, and Y. He, "Nucleus segmentation of cervical cytology images based on multi-scale fuzzy clustering algorithm," *Bioengineered*, vol. 11, no. 1, pp. 484–501, 2020.
- [29] K. H. S. Allehaibi, L. E. Nugroho, and L. Lazuardi, "Segmentation and classification of cervical cells using deep learning[J]," *IEEE Access*, vol. 7, pp. 116925–116941, 2019.
- [30] Y. Song, E. L. Tan, X. Jiang et al., "Accurate cervical cell segmentation from overlapping clumps in pap smear images," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 288–300, 2017.
- [31] C. Chunming Li, C. Chenyang Xu, C. Changfeng Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3243–3254, 2010.
- [32] A. Sakthi and M. Rajaram, "Density based multiclass support vector machine using IoT driven service oriented architecture for predicting cervical cancer," *International Journal of u and e-Service, Science and Technology, SERSC Australia*, vol. 9, no. 11, pp. 195–216, 2016.
- [33] T. Guan, D. Zhou, and Y. Liu, "Accurate segmentation of partially overlapping cervical cells based on dynamic sparse contour searching and GVF snake model," *IEEE journal of biomedical and health informatics*, vol. 19, no. 4, pp. 1494–1504, 2015.
- [34] A. Ray, I. Kanta Maitra, and D. Bhattacharyya, "Detection of cervical cancer at an early stage using hybrid segmentation techniques from PAP smear images," *International Journal of Advanced Science and Technology, SERSC Australia*, vol. 112, pp. 23–32, 2018.
- [35] P. Soille, *Morphological Image Analysis: Principles and applications[M]*, Springer Science & Business Media, Berlin, Heidelberg, 2013.
- [36] J. E. Bresenham, "Algorithm for computer control of a digital plotter," *IBM Systems Journal*, vol. 4, no. 1, pp. 25–30, 1965.
- [37] A. Tareef, Y. Song, H. Huang et al., "Multi-pass fast watershed for accurate segmentation of overlapping cervical cells," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2044–2059, 2018.
- [38] M. Nosrati and G. Hamarneh, "A variational approach for overlapping cell segmentation[J]," *ISBI overlapping cervical cytology image segmentation challenge*, pp. 1–2, 2014.

- [39] G. L. B. Ramalho, D. S. Ferreira, and A. G. C. Bianchi, "Cell reconstruction under voronoi and enclosing ellipses from 3D microscopy[C]," in *Proceedings of the Second Overlapping Cervical Cytology Image Segmentation Challenge-IEEE International Symposium on Biomedical Imaging (ISBI 2015)*, IEEE, Rio Grande, Brazil, October 2015.
- [40] T. Wang, "Overlapping cell segmentation with depth information[C]," in *Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications*, pp. 278–287, Springer, Cairo, Egypt, May 2022.