Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 20(9): 16259–16278.
DOI: 10.3934/mbe.2023726
Received: 18 May 2023
Revised: 07 August 2023
Accepted: 08 August 2023
Published: 14 August 2023

Research article

Cell segmentation in fluorescence microscopy images based on multi-scale histogram thresholding

Yating Fang and Baojiang Zhong*

School of Computer Science and Technology, Soochow University, Suzhou 215021, China

* **Correspondence:** Email: bjzhong@suda.edu.cn.

Abstract: Cell segmentation from fluorescent microscopy images plays an important role in various applications, such as disease mechanism assessment and drug discovery research. Exiting segmentation methods often adopt image binarization as the first step, through which the foreground cell is separated from the background so that the subsequent processing steps can be greatly facilitated. To pursue this goal, a histogram thresholding can be performed on the input image, which first applies a Gaussian smoothing to suppress the jaggedness of the histogram curve and then exploits Rosin's method to determine a threshold for conducting image binarization. However, an inappropriate amount of smoothing could lead to the inaccurate segmentation of cells. To address this crucial problem, a multi-scale histogram thresholding (MHT) technique is proposed in the present paper, where the scale refers to the standard deviation of the Gaussian that determines the amount of smoothing. To be specific, the image histogram is smoothed at three chosen scales first, and then the smoothed histogram curves are fused to conduct image binarization via thresholding. To further improve the segmentation accuracy and overcome the difficulty of extracting overlapping cells, our proposed MHT technique is incorporated into a multi-scale cell segmentation framework, in which a region-based ellipse fitting technique is adopted to identify overlapping cells. Extensive experimental results obtained on benchmark datasets show that the new method can deliver superior performance compared to the current state-of-the-arts.

Keywords: cell segmentation; multi-scale; histogram thresholding; ellipse fitting; fluorescence microscopy images

1. Introduction

Numerous areas of research on analyzing and quantifying fluorescence microscopy images rely on quantitative cell nucleus image analysis [1]. In particular, accurate cell segmentation in fluorescent microscope images plays a key role in high-throughput applications such as protein expression quantification and cell function research [2]. However, manual or semi-automatic cell segmentation is a

time-consuming task. Noise and the overlapping of adjacent cells make this task even more difficult. Thus, automatic cell segmentation using image processing techniques has attracted considerable interest [3–5].

Over the past few decades, a great number of methods have been proposed for cell segmentation. Mueller et al. [6] developed a segmentation method based on sparse component analysis to separate cell nuclei, fibrous components and adipose components. Zhao et al. [7] made use of channel-fused images with seed-based cluster splitting and the k-means algorithm to segment individual nuclei with fuzzy boundaries or in a clump. Riccio et al. [2] combined distance transform and curvature analysis to separate overlapped cells. One of the most common approaches for cell segmentation is to exploit morphological operations. Wang and Wang [8] proposed a method for segmenting a variety of cells robustly; it uses the Gabor filter to enhance the gradient image, which is followed by segmentation of the enhanced gradient image via threshold selection based on the slope difference distribution. Observing that over-segmentation could be frequently yielded when cells are clustered, the watershed transform-based method was proposed. Salvi et al. [9] proposed a watershed-based segmentation method to segment human cardiac cells in fluorescence microscopy images. Jia et al. [10] combined the marker-controlled watershed algorithm with a two-step method based on the watershed, split and merge watershed, which can effectively decrease the level of over-segmentation.

Another promising class of methods performs image binarization as the first step, which separates foreground cells from the background so that the subsequent processing steps of cell segmentation can be greatly facilitated [11, 12]. The image binarization step can be done by thresholding the image histogram [13, 14]. The basic idea of histogram thresholding is to determine an appropriate threshold for satisfactory image binarization by analyzing the histogram of the image. A closely related procedure is multi-level thresholding image segmentation. Ewees et al. [15] developed a hybrid metaheuristic approach for efficient multi-level image segmentation by integrating both the artificial bee colony algorithm and the sine-cosine algorithm. Yang et al. [16] used the metaheuristic method and multi-level threshold image segmentation technology to segment breast cancer pathological images and provided outstanding samples for the subsequent analysis of breast cancer computer-aided diagnosis. In these studies, the 2D histogram thresholding has been suggested to improve the efficiency of subsequent image analysis and diagnosis.

Deep learning methods have also been developed for cell image segmentation. Raza et al. [17] proposed a multiple-input multiple-output convolutional neural network, which integrates the features of pyramidal images into the U-Net for the purpose of applying cell segmentation to fluorescence microscopy images. Wang et al. [18] introduced a channel transformer network for reliable and automatic segmentation of medical images. Cao et al. [19] proposed the Swin-Unet, which is a U-Net-like pure transformer for medical image segmentation. In general, a deep learning model can achieve high performance if it has been adequately trained on large-scale datasets. In the field of medical image segmentation, however, the number of data samples is often relatively small, making it rather difficult to sufficiently train the deep learning model for specific applications, e.g., our considered cell segmentation task for fluorescent microscopy images. In such a case, non-deep learning methods, including the one we have developed in this work, provide a more suitable solution since they do not require any pre-training. Moreover, our proposed method has the advantages of broad applicability (e.g., being freely used to process both 8-bit and 16-bit images), high computational efficiency and easy implementation. Thus, it is suitable for a wide range of embedded systems where GPU resources are unavailable.

Ellipse fitting is gradually becoming a popular approach for facilitating cell segmentation. Prasad et al. [20] proposed a least squares-based ellipse fitting technique, which has high selectivity for elliptic curves. However, it could exhibit poor performance in the presence of inflection points on edge contour arcs. Following the work of Prasad et al. [20], Shen et al. [21] encoded arcs and their connectivity into a sparse directed graph to accelerate the process of arcs pairing. Zafari et al. [22] developed an ellipse fitting method that can separate overlapping objects in silhouette images. However, this method might exhibit poor performance in the presence of non-Gaussian outliers. Das et al. [23] proposed an effective hybrid ellipse fitting method for blood-cell segmentation. However, we observed that it suffers from the inability to detect elongated cells with high accuracy, since the seed points of such blood cells could have large location error.

To address the above-mentioned problems and improve the cell segmentation performance, a novel multi-scale histogram thresholding (MHT) technique is proposed, which is then further incorporated into a multi-scale framework for conducting cell segmentation in this study. Our motivation to adopt the multi-scale methodology originates from the fact that insignificant local structures in and noise effects on a given signal can be effectively suppressed by using the Gaussian filter over a set of scales, where the scale refers to the standard deviation of the Gaussian. Note that, in the literature, there are a few so-called "multi-scale" cell segmentation methods; however, they commonly adopt a multi-scale operation within one step of cell segmentation only. For example, Ram and Rodriguez [14] used the multi-scale variance stabilizing transform as a preprocessing step for noise suppression. Al-Kofahi et al. [24] used the multi-scale Laplacian-of-Gaussian filtering in the step of seed detection. In comparison, our developed cell segmentation method is a true or full-blown multi-scale approach, as it has multiple branches that perform all cell segmentation steps individually over multiple scales. The main contributions of our work are summarized as follows.

- Propose an MHT approach for separating foreground cells from the background, which overcomes the difficulty of the existing single-scale approach in determining an appropriate scale that can work in different real-world circumstances.
- Identify overlapping cells with a region-based ellipse fitting technique, which addresses the challenge of extracting overlapping cells from the input image correctly.
- Develop a multi-scale framework for cell segmentation in fluorescence microscopy images, which performs cell segmentation over a set of scales to improve the segmentation accuracy.

The remainder of the paper is organized as follows. The MHT technique is proposed in Section 2. Our multi-scale cell segmentation framework is described and investigated in Section 3. In Section 4, experimental results are provided and discussed. Finally, this paper is concluded in Section 5.

2. Multi-scale histogram thresholding

2.1. Motivation

Let $H(x, y)$ be the histogram of the input fluorescence microscopy image, where $x \in [0, 255]$ denotes the gray level of pixels and y denotes the frequency of x. To conduct a multi-scale analysis, the histogram is convolved with a 2D Gaussian kernel g_σ with varying standard deviations, σ, as follows:

$$H_\sigma(x, y) = (H * g_\sigma)(x, y), \tag{2.1}$$

where the symbol $*$ denotes the convolution operator, σ serves as the scale parameter and $H_\sigma(x, y)$ is the smoothed histogram.

We employ the Gaussian filtering to execute histogram smoothing, which can effectively reduce aliasing artifacts. However, the single-scale method may not be optimal for smoothing these artifacts, as the histogram can become excessively smoothed with increasing scale, leading to a decrease in image contrast. Therefore, we combine the histogram frequency information on a set of scales (i.e., high, middle and low scales) and average them as the final result to overcome the drawbacks of the single-scale method. The MHT technique is thus developed through this frequency averaging strategy, detailed as follows.

Figure 1 depicts the underlying idea of our proposed MHT technique. The test image shown in Figure 1(a) has been cropped from a fluorescence microscopy image, which contains a single cell. The challenge to segmenting this cell from the background lies in that there are many pixels inside the cell whose intensity is close to the background, resulting in the histogram not being smooth enough, and the selected threshold is not optimal. The choice of threshold is explained in detail in Section 2.2.4. Figure 1(f) is a smooth histogram at a low scale. It can be seen that the bottom-left corner of the histogram is not smoothed enough; thus, the selected threshold is too small, which causes the background to be segmented as the foreground, as shown in Figure 1(c). Figure 1(g) is a smoothed histogram at a high scale. It can be seen that the histogram is oversmoothed, so the selected threshold is too large, which results in parts of a cell being filtered out as noise, as shown in Figure 1(d). In our approach, a set of scales is used, and the multi-scale result is obtained by fusing the histogram curve information on these scales and taking the average result as output, as shown in Figure 1(h). The determined threshold can accurately segment the foreground cells without mis-segmenting the background into the foreground, as shown in Figure 1(e).

2.2. Methodology

The outline of our proposed MHT technique is shown in Figure 2. First, a histogram of the fluorescence microscopy image is obtained by counting the pixel frequency. Next, histogram smoothing is executed over three chosen scales. A multi-scale fusion is then performed to produce a smoothed histogram H_f. Finally, the threshold T to segment the foreground cells from the background is determined on H_f by using Rosin's method [25]. In what follows, we will start from a brief introduction to the histogram to describe our multi-scale thresholding technique.

2.2.1. Histogram of the image

The histogram, also known as the frequency histogram, is used to show the distribution of a data set. The histogram of an image is used to characterize the distribution of pixel values of the image, i.e., to display the frequency of occurrence of different pixel values at different intensity values. For gray-scale images, the intensity range is between $[0, 255]$, and for RGB color images, the image histograms of three colors can be independently displayed. In this paper, we use the histogram of the gray-scale image, whose pixel frequency counting process is as follows:

$$H_I = \mathrm{FC}(I_s, F_j), \tag{2.2}$$

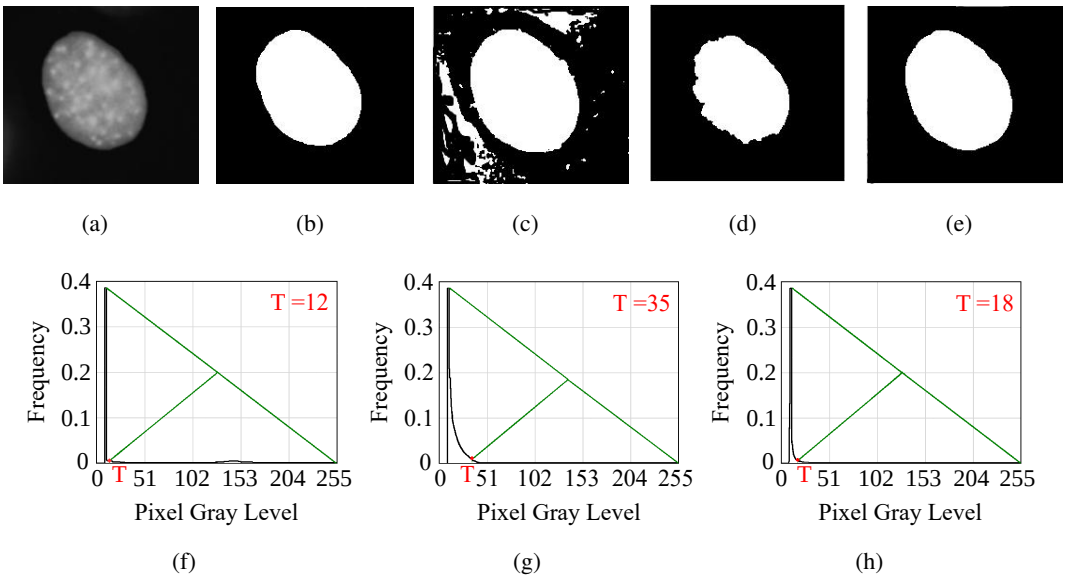


Figure 1. Underlying idea of our proposed multi-scale histogram thresholding (MHT) technique: (a) Test image; (b) Ground truth; (c)–(e) Image binarization results for the existing single-scale histogram thresholding performed at a low scale and a high scale, as well as those for our MHT, respectively; (f)–(h) Smoothed histogram curves for the test image obtained at the low and high scales, as well as that yielded via our MHT through a multi-scale fusion, respectively.

with

$$F_j = \frac{1}{N}\sum_{j=0}^{255}(p_i == j), \tag{2.3}$$

where H_I represents the histogram of the input image I_s, F_j represents the frequency of the pixel with the gray-scale value j ($j = 0, 1, ..., 255$) on I_s, p_i is the i-th pixel ($i = 1, 2, \cdots, N$) on the image I_s, the condition ($p_i == j$) means that 'if p_i equals j' and N denotes the total number of pixels.

2.2.2. Individual histogram smoothing over multiple scales

In this work, three scales are used: $\sigma_1 = 5$, $\sigma_2 = 10$ and $\sigma_3 = 20$, respectively. The Gaussian smoothing executed on the image histogram H_I over the three specified scales can be expressed as:

$$H_I{}^{(k)} = H_I * g_{\sigma_k}, \tag{2.4}$$

where $H_I{}^{(k)}$ denotes the smoothed histogram yielded at the k-th scale σ_k for $k = 1, 2, 3$, respectively.

2.2.3. Multi-scale fusion

Assume that the smoothed histogram $H_I{}^{(k)}$ yielded at the k-th scale consists of a set of points as $P^{(k)} = \left\{p_i^{(k)}\right\}_{i=1}^{256}$, where $p_i^{(k)}$ represents the coordinates of each point on the smoothed histogram curve. To pursue the goal of MHT, the final multi-scale point set P_f is produced by performing frequency averaging calculation as follows:

$$P_f = \frac{1}{3}\sum_{k=1}^{3} P^{(k)}. \tag{2.5}$$

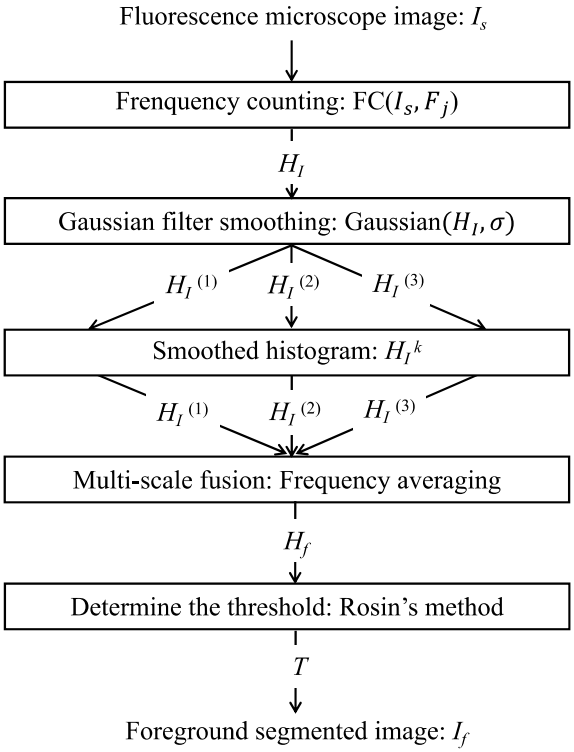


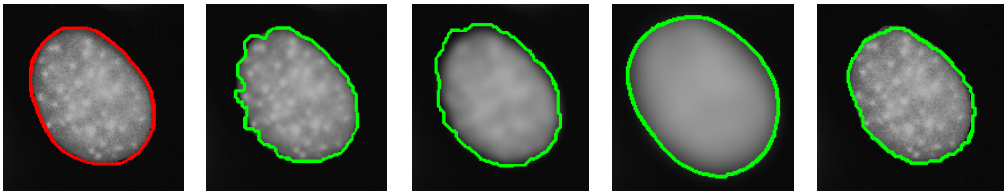
Figure 2. Pipeline of our proposed MHT technique.

As a result, the final histogram H_f is a smooth curve obtained by connecting all of the points in P_f.

2.2.4. Threshold selection

The basic idea of the histogram thresholding technique is to determine a threshold on the gray-scale values of pixels to thus separate the foreground cell from the background. After performing the above-mentioned multi-scale fusion to remove the influence of illumination and noise, the histogram curve exhibits double peaks, one related to the background and the other related to the foreground object. Note that such a histogram curve is often said to be unimodal since the peak corresponding to the foreground cell is generally very inconspicuous. The valley between the two peaks takes the shape of an "L", where the segmentation threshold should be selected. Based on the L-curve criterion [26, 27], we take the "corner" of the L-shape as the most appropriate choice for the segmentation threshold, denoted as T. In fact, the chosen T can achieve the best trade-off in the threshold selection, as a smaller threshold will identify more of the background region as the cell foreground, while a larger threshold will take more of the cell region as the background.

Rosin's method [25] is then used to find the corner of the L-shape, i.e., the segmentation threshold T is finally determined, as shown in Figure 1(h). First, the highest and lowest points on the histogram are identified and a straight line is drawn to connect them. After that, the corner of the L-shape is recognized as the point on the histogram curve H_f that is farthest from this straight line. Finally, the gray-scale value corresponding to the horizontal coordinate value of the corner point is chosen as the threshold T. Compared with gradient-based methods and stochastic optimization-based methods, Rosin's method [25] is computationally more efficient, easier to implement and demonstrates broader



GT / IoU (a) | Low / 0.84 (b) | Medium / 0.93 (c) | High / 0.90 (d) | Ours / 0.96 (e)

Figure 3. Performance of our developed multi-scale cell segmentation method: (a) Ground truth; (b)–(d) Cell segmentation performed at the low, medium and high scales, respectively; (e) The segmentation result of our developed method, produced by averaging the outcomes of the three scales presented in (b)–(d).

applicability across different images.

3. Materials and methods

Existing cell segmentation methods generally extract cell contours at a single scale only, where the scale refers to the standard deviation of the Gaussian filter applied to smooth the input image and thus suppress insignificant local image structures. However, if the scale value is set too small, then the smoothing effect is insufficient. As a result, local image structures could be incorrectly taken as part of the cell contours, as shown in Figure 3(b). On the other hand, if the scale value is chosen too large, then the cell contours will spread outward due to image blurring, leading to poor segmentation accuracy, as demonstrated in Figure 3(d). Note that, although the medium scale seems to yield a promising cell contour extraction result in our illustrative example, it is hard to determine a fixed scale that can be applicable to all input images in real-world applications.

To overcome the above-mentioned dilemma, our proposed MHT technique is further incorporated into a multi-scale cell segmentation framework. The developed multi-scale method is outlined in Figure 4. It has three branches, each of which extracts the cell contours at a chosen scale, followed by fusing all of the obtained cell contours to arrive at the final segmentation result. From a signal processing point of view, this multi-scale fusion represents a signal averaging operation applied to the scale domain, which can greatly increase the strength of the signal relative to the noise [28]. In our case, the position accuracy of cell contours is thus improved relative to that of the existing single-scale methods. The effect of our proposed multi-scale method is depicted in Figure 3(e). One can see that a satisfactory segmentation result is produced by averaging the cell contours extracted over the three used scales.

We assume a gray-scale image (see Figure 6(a)) containing a number of cells that may vary in size and shape and may be touching each other. The cells may vary in brightness. This might also apply to the background of the cell. Each cell is free of holes and can be distinguished from its local background because of its higher brightness and its elliptic-like shape. First, the processes of image smoothing, segmenting cells from their background and identifying overlapping cells are sequentially conducted over multiple scales. A multi-scale fusion is then performed to produce the cell segmentation result. In what follows, we will start from the image-smoothing step to describe our multi-scale cell segmentation method.

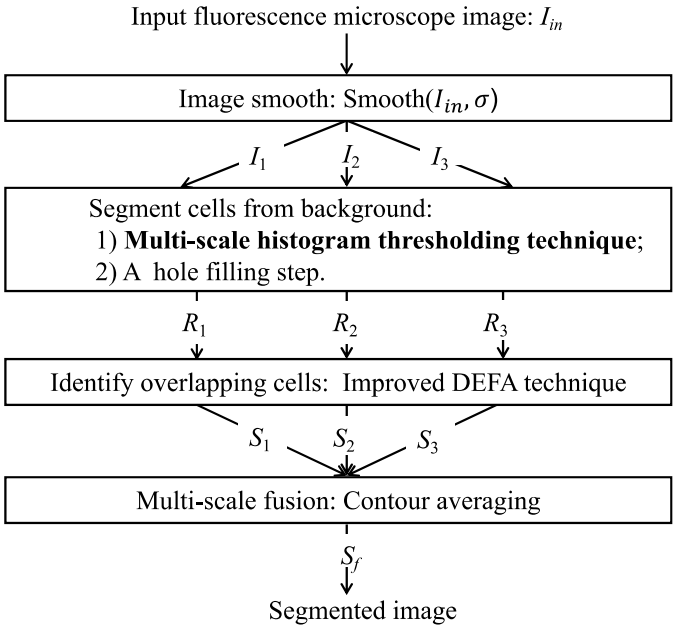


Figure 4. Outline of our proposed multi-scale cell segmentation method.

3.1. Individual image smoothing over multiple scales

In the digital case, the image convolution process is performed with a discrete sequence of scales. In this work, three scales are used: $\sigma_1 = 2/\sqrt{2}$, $\sigma_2 = 2$ and $\sigma_3 = 2\sqrt{2}$, respectively. Note that the medium scale σ_2 is equal to the default scale of the existing single-scale method [29]. The use of high scale σ_3 helps to smooth the contour of the cell segmentation, and the use of low scale σ_1 helps to improve cell location accuracy.

Assume that the input fluorescence microscopy image is denoted as I_{in}. The image-smoothing operation over the three specified scales can be expressed as

$$I_i = I_{in} * g_{\sigma_i}, \tag{3.1}$$

where I_i denotes the Gaussian blurred image yielded at the i-th scale by using the Gaussian kernel g_{σ_i} for $i = 1, 2, 3$, respectively.

3.2. Segmenting cells from background

Fluorescence microscopy cell images contain many cells of various sizes with elliptic-like shapes, which may be touching each other. The brightness of these cells differs greatly from their background. Combining this brightness difference between foreground cells and background, we apply the MHT technique and a hole-filling step to segment cells from their background. The MHT technique has been detailed in Section 2.

There are a few regions in the highly textured cells that are black and are hence detected as background, as shown in Figure 5(b). To avoid this, we employ a hole-filling step to fill these black regions. Figure 5(c) shows the result of filling all holes in the cell.

In addition, a drawback of the histogram thresholding technique is that segments of the background with locally higher brightness are identified erroneously as cells (see Figure 6(b)). To reduce these false

(a) (b) (c)

Figure 5. Effect of the hole-filling step. (a) Test image; (b) Image binarization result, where a few holes are seen in the cell region; (c) Outcome of hole filling.

positives, we have introduced two shape-based constraints: area constraint and roundness constraint, as well as two intensity-based constraints: local and global intensity constraints.

Area constraint: Assume an expected area S_e of each cell in images and a minimum area threshold S_α. The area constraint can be defined as $S_e >= S_\alpha$, which can avoid particularly small background segments being considered as cells. In addition, to prevent partially visible cells, which appear at the image boundary, from being rejected as such background segments, the expected area S_e should be calculated as the area S_c of the circle that can be fitted best to the eight extrema points of their boundary (as shown in Figure 6(c)) [29], rather than its real area S_r.

Roundness constraint: The roundness R can measure the similarity between the shape of an object and a standard circle, which is defined as

$$R = 4\pi S/C^2, \tag{3.2}$$

where C and S represent the perimeter and area of the object, respectively. Since the cells are mostly circular/elliptic-like, the roundness constraint is to use the roundness measure R to reject objects that deviate from this shape. According to our experiments, $R > 0.2$ is required.

Local intensity constraint: The shape constraints can reject several false positives, such as the three on the right in Figure 6(b). To reject more false positives, for example, in the lower-left and upper-left corners of Figure 6(b), we introduce the strength constraints. In the local intensity constraint, the intensity distribution within a cell should be more similar to the intensity distribution within the rest of the cells, rather than to the intensity distribution of the local background. It is effective in rejecting cell-shape false positives, such as the circular object at the top left of Figure 6(b). Specifically, we first extract the local background of each detected object by computing the Voronoi diagram of the object centroid; we then remove the detected object from it (see Figure 6(c)). Then, we adopt the Bhattacharyya distance [29] to measure the distance between two intensity distributions. Assume that the Bhattacharyya distances between the detected object and the rest of the objects and the local background are d_0 and d_1, respectively; the detected object is rejected if $d_0 > d_1$.

Global intensity constraint: There will be strong-brightness, ellipse-like background segments in fluorescence microscopy images, with similar intensity cells nearby. In this case, they will still be mistaken for cells under the local intensity constraint (as shown in Figure 6(e) lower-left corner). To address this issue, we introduce a global intensity constraint. This constraint can be expressed intuitively by using OTSU [30] to determine the image's global threshold level, and by resetting the segment whose intensity value is greater than the level in the image to the level.

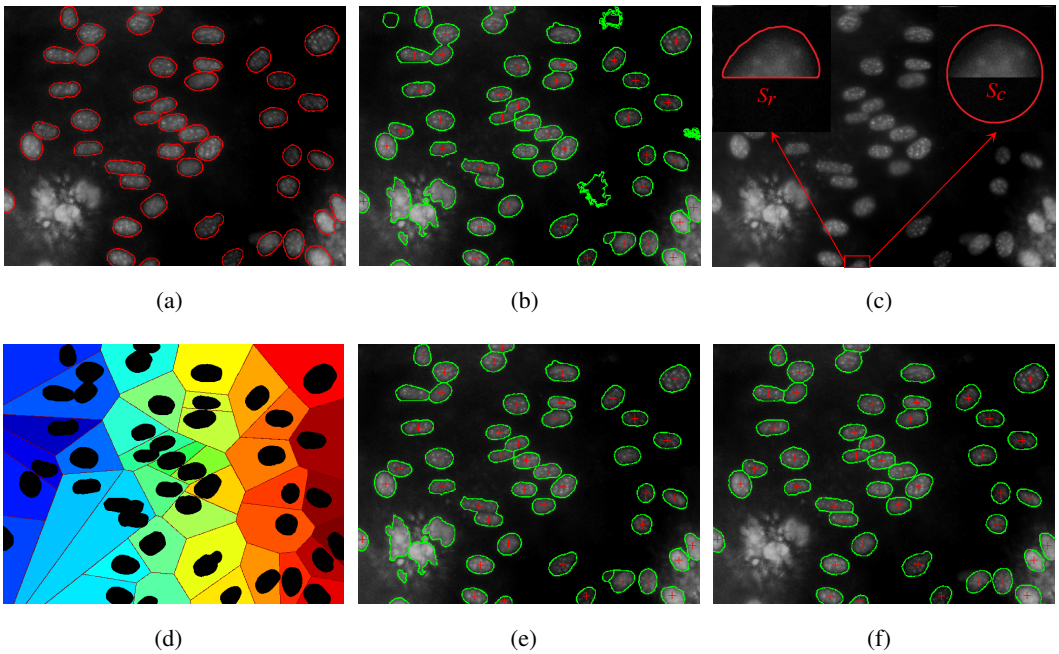


Figure 6. Segmenting cells from the background. (a) Ground truth; (b) Cell contours yielded by using our proposed MHT, where the centroid of each ground-truth cell is marked with a "+" for verification; (c) Effect of the area constraint on segmenting an occluded cell (with magnified illustration), where S_c and S_r are the measured areas of the cell with and without using this constraint, respectively; (d) The local backgrounds of the detected cells (in black) according to the Voronoi diagram of their centroids; (e) Segmentation result of the RFOVE method; (f) Segmentation result of our proposed method.

Figure 6(b) shows the boundaries of the detected cells by the original histogram thresholding technique, and the ground truth is shown in Figure 6(a). These false positives are rejected by employing the proposed constraints (see Figure 6(f)).

3.3. Identifying overlapping cells

With the separation of the foreground cells from the background, a set of connected regions are obtained, denoted as $R = \{r_1, r_2, \dots\}$. In what follows, the decreasing ellipse fitting algorithm (DEFA) [31] is employed to identify whether each region corresponds to overlapping cells.

The DEFA fits an arbitrary 2D shape with a number of ellipses to count the number of objects that make up the shape. In the DEFA, the number of ellipses and their parameters are determined automatically under the constraint that the total area $|E| = \sum_{i=1}^{k} |E_i|$ of the ellipses is equal to the area A of the concerned shape. When incorporating the DEFA into our cell segmentation process, a skeleton of the current region r_j is computed first, which provides information about the fitted ellipse parameters of r_j. Then, different models (i.e., solutions involving different numbers of ellipses) are evaluated based on the Akaike information criterion (AIC) [29]. Specifically, from all possible models, the model with the smallest AIC is reported as the optimal solution. To minimize the AIC, we start with a custom set of initial size $|CC|$ and gradually reduce the number of considered ellipses until the set of all ellipses

contains a single ellipse. In each iteration, a pair of adjacent ellipses is selected for merging. The pair finally merged is the one that results in the lowest AIC. Finally, the cell pixels are clustered into groups according to the detected ellipses, while keeping the region boundaries previously obtained from the MHT to complete the segmentation of overlapping cells. The generated cell contours are denoted as $S = \{l_1, l_2, \dots\}$, where l_i represents the contour of the i-th cell and $i = 1, 2, \cdots$.

3.4. Multi-scale fusion

To cope with cell features at different scales, the previous steps of segmenting cells from their background and identifying overlapping cells steps are performed over three scales for the purpose of conducting multi-scale cell segmentation. Denote a feature point of the i-th cell contour at the j-th scale as $p_{i_k}^{(j)}$, where $i, k = 1, 2, ...$ and $j = 1, 2, 3$. A multi-scale feature point $\bar{p}_{i_k}$ of the i-th cell contour is produced by performing the signal averaging calculation as follows:

$$\bar{p}_{i_k} = \sqrt[3]{\prod_{j=1}^{3} p_{i_k}^{(j)}}. \tag{3.3}$$

As a result, the final cell contour set S_f is obtained by combining all feature points, i.e.,

$$S_f = \left\{\bar{l}_i : \bar{l}_i = \{\bar{p}_{i_k}\}, i, k = 1, 2, \dots\right\}, \tag{3.4}$$

where $\bar{l}_i$ denotes the final i-th cell contour. In this way, each cell is segmented from the input image.

Specifically, cell contours appearing at the high scale will also appear at the low and middle scales since it never introduces new features into the smoothed image [32]. In addition, the resulting cell contours at low scales are not smooth enough due to noise. As the scale increases, the noise influence gradually decreases and the resulting cell contours at middle and high scales become closer to the true shape of the cell. However, the increase in scale can lead to less accurate positioning of cells. Therefore, the cell contour set yielded at each scale σ_j, i.e., S_j, is taken as the initial set of cell contour candidates. Then, the image features at each scale, that is, the cell contour, are fused, and the signal averaging result S_f is used as the final output result, thereby segmenting each cell from the input image.

4. Results

4.1. Experimental setups

4.1.1. Datasets

To validate our proposed method, two benchmark datasets [33] were used: the NIH3T3 and the U20S. These two datasets consist of 49 and 48 fluorescence microscopy cell images (1344 × 1024 pixels), respectively. The NIH3T3 is considered to be more challenging since its images contain cells/nuclei that vary greatly in brightness, as well as visible debris [33]. The source code for our proposed method has been made publicly available at https://github.com/Aurora-34 0f/MHT.git.

4.1.2. Objective evaluation metrics

For objective evaluation, we employed both region-based and contour-based metrics to evaluate the performance of the cell segmentation, as practiced in [1] and [34]. The region-based metrics include

the Jaccard coefficient (JSC), aggregated Jaccard index (AJI), Dice false positive (DiceFP) and Dice false negative (DiceFN). As contour-based metrics, we used the Hausdorff distance (HD) and mean absolute contour distance (MAD). In addition, the numbers of false positives and false negatives have been used to evaluate the cell splitting performance.

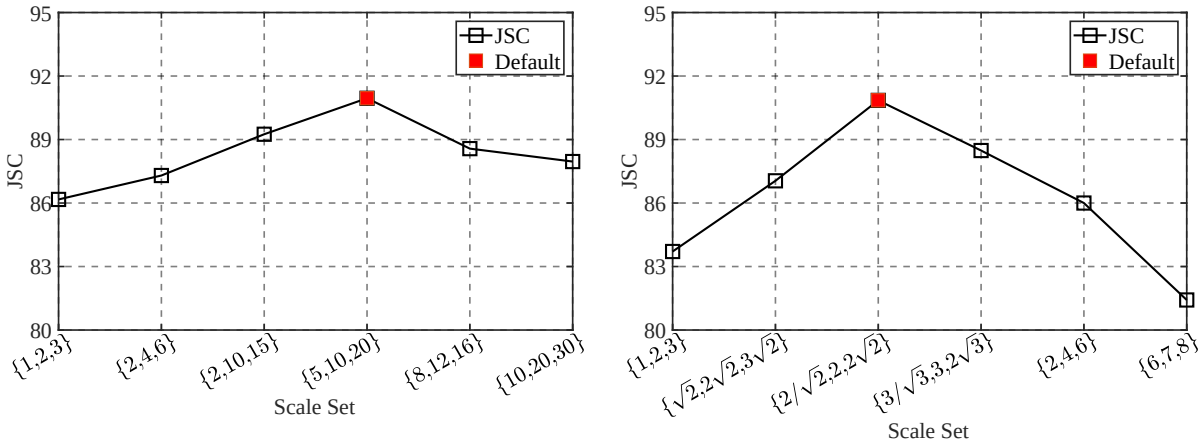


Figure 7. Cell segmentation performance in terms of the JSC, plotted against different scale settings. (a) Our proposed MHT technique; (b) Our developed multi-scale cell segmentation method. Note that a higher JSC indicates better performance.

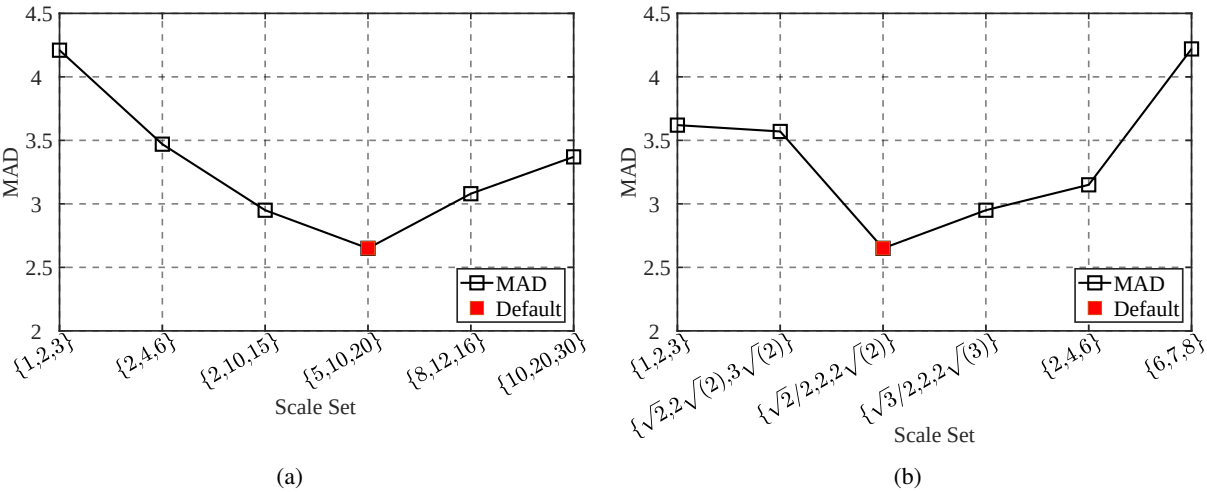


Figure 8. Cell segmentation performance in terms of the MAD, plotted against different scale settings. (a) Our proposed MHT technique; (b) Our developed multi-scale cell segmentation method. Note that a lower MAD indicates better performance.

Two datasets, U20S and NIH3T3, were used to tune the scaling parameters (i.e., the values of σ_1, σ_2 and σ_3). Our proposed MHT and multi-scale cell segmentation methods were implemented on image sets with different scales. The cell segmentation performances measured according to the JSC

(area-based metric) and MAD (contour-based metric) are respectively shown in Figures 7 and 8. The set of scales that produced the highest JSC and lowest MAD was taken as the default setting for the scale parameters of our proposed MHT and multi-scale cell segmentation techniques. Specifically, the scale parameters of the proposed MHT technique were $\sigma_1 = 5$, $\sigma_2 = 10$ and $\sigma_3 = 20$ respectively, and the scale parameters of the multi-scale cell segmentation method were $\sigma_1 = 2/\sqrt{2}$, $\sigma_2 = 2$ and $\sigma_3 = 2\sqrt{2}$.

4.1.3. Scale parameter tuning

Table 1. Performance comparison for different methods based on five commonly used evaluation metrics, where ↑ (↓) indicates that the higher (lower) the used evaluation metric, the better the segmentation performance. The best performance in each comparison case is highlighted in bold.

Dataset	Method	JSC ↑	AJI ↑	MAD ↓	HD ↓	DiceFN ↓	DiceFP ↓
U20S	LSBR [36]	83.2	-	5.8	19.8	9.1	11.8
	Three-step [35]	88.4	-	4.7	13.4	5.2	5.3
	LLBWIP [1]	91.6	-	3.5	12.7	3.9	4.7
	OTSU [30]	83.5	57.6	4.5	11.5	16.7	3.0
	UCTransNet [18]	80.7	55.8	6.2	20.3	9.4	11.9
	Swin-Unet [19]	80.3	56.2	6.6	21.7	10.6	11.2
	RFOVE [29]	89.8	74.9	2.8	7.5	5.7	5.2
	Proposed	**94.1**	**76.6**	**2.0**	**5.7**	**3.2**	**2.9**
NIH3T3	LSBR [36]	64.2	-	7.2	19.8	20.4	21.2
	Three-step [35]	70.8	-	5.7	16.4	19.7	15.5
	LLBWIP [1]	75.9	-	4.1	14.3	12.2	12.7
	OTSU [30]	59.6	24.0	6.2	12.9	35.4	24.2
	UCTransNet [18]	73.8	47.7	5.1	14.8	14.5	15.6
	Swin-Unet [19]	69.9	46.4	6.4	17.2	20.3	15.2
	RFOVE [29]	81.0	63.9	3.5	**8.2**	8.1	13.3
	Proposed	**87.8**	**70.5**	**3.3**	**8.2**	**5.1**	**7.9**

Table 2. Performance comparison for different methods based on the false posotive and false negative rates. The best performance in each comparison case is highlighted in bold.

Methods	U20S		NIH3T3	
	FP	FN	FP	FN
Three-step [35]	0.5	3.9	1.7	11.3
LLBWIP [1]	0.3	2.7	0.7	0.8
UCTransNet [18]	1.3	3.3	1.6	4.1
Swin-Unet [19]	2.0	3.9	2.7	5.9
MSF [10]	1.9	4.6	1.8	1.7
RFOVE [29]	1.9	0.3	**0.3**	0.8
Proposed	**0.2**	**0.1**	**0.2**	**0.5**

4.2. Objective evaluation

4.2.1. Comparison with non-learning-based existing methods

In total, six state-of-the-art non-learning-based methods have been used for comparison, including the OTSU method with a hole-filling step, as described in [30], as well as the three-step [35], LSBR [36], LLBWIP [1], MSF [10] and RFOVE [29] methods. Table 1 summarizes the segmentation results obtained on the U20S and NIH3T3 datasets respectively. It can be seen that the proposed method produces the best performance for all of the used evaluation metrics on both the U20S dataset and the NIH3T3 dataset. Note that, since the LSBR [36], three-step [35] and LLBWIP [1] methods have neither publicly available codes nor publicly available evaluation data for the AJI performance, they have not been evaluated in terms of this evaluation metric. Table 2 gives the evaluation of splitting results on the U20S and NIH3T3 datasets. As can be seen from Table 2, the proposed method works best on the division of U20S and NIH3T3 cells.

4.2.2. Comparison with deep learning-based methods

Furthermore, the proposed method is compared with two deeplearning-based methods, i.e., UCTransNet [18] and Swin-Unet [19], to demonstrate the proposed method's superiority. According to [18], we divided the 49 images of NIH3T3 into three sets: 29 for training, six for validation and 14 for testing, and the 48 U20S images into three sets: 28 for training, six for validation and 14 for testing. As can be seen from Table 1 and Table 2, our proposed method results are significantly better than the results of the UCTransNet [18] and the Swin-Unet [19].

4.3. Subjective evaluation

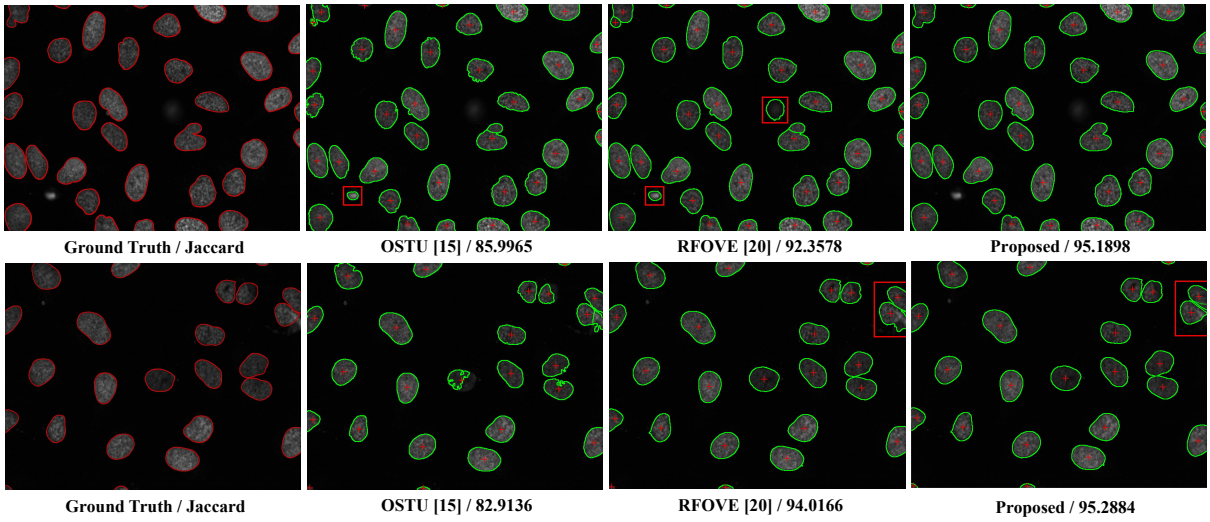


Figure 9. Subjective comparison of different cell segmentation methods based on two test images taken from the U20S dataset. The contours of detected cells are shown by green curves, and, for reference, the centroids of ground-truth cells are marked with red "+".

Figure 9 shows a subjective comparison of our proposed and the existing methods based on test images from the U20S dataset. In the first row, the test image contains locally brighter background

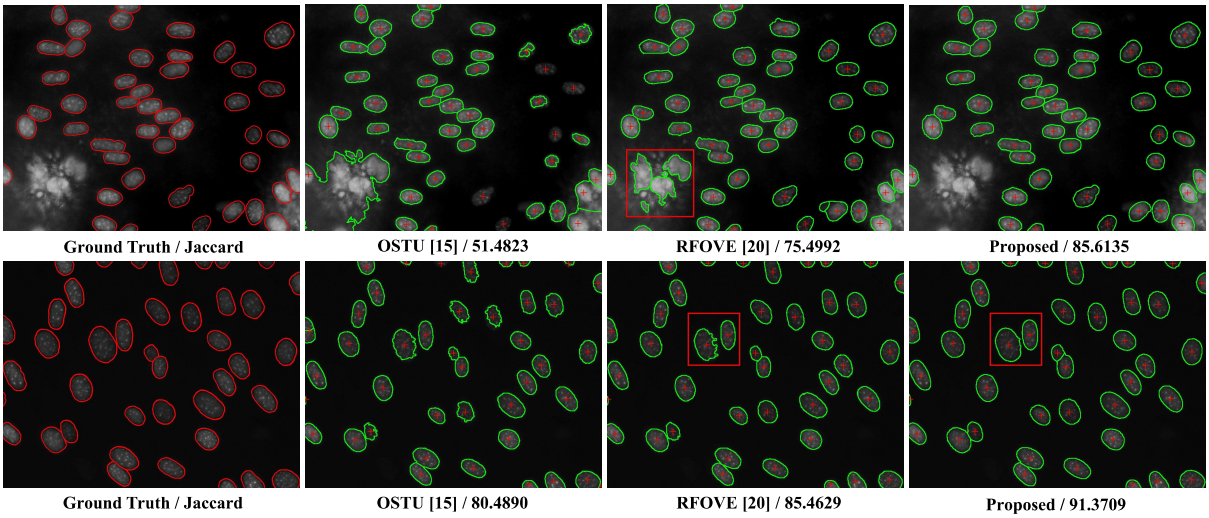


Figure 10. Subjective comparison of different cell segmentation methods based on two test images taken from the NIH3T3 dataset. The contours of detected cells are shown by green curves, and, for reference, the centroids of ground-truth cells are marked with red "+".

fragments, which RFOVE and OTSU misidentify as cells and yield false positives (such as the part circled by the red frame), while our proposed method can accurately reject them. In the second row, due to the influence of noise, cell edges generated by RFOVE segmentation are not smooth enough (such as the part circled by the red frame); most of the cell edges generated by OTSU segmentation are not smooth enough. However, our proposed method could effectively reduce the influence of noise and produce smoother and more accurate cell edges.

Figure 10 shows a subjective comparison of our proposed and the existing methods based on test images from the NIH3T3 dataset, whose result analysis is similar to that of Figure 9. It is worth reminding that the NIH3T3 dataset is more challenging, containing cells and background segments with widely varying brightness, which makes it difficult to segment cells by using a single global threshold. Therefore, the segmentation performance of OTSU in Figure 10 is relatively poor. For example, in the first row, a large number of cells were not accurately segmented by OTSU, resulting in many false negative results; in the second row, OTSU segmented out lots of cells with rough contours. In comparison, our proposed method delivered the best subjective results with smooth and accurate segmented cell contours.

4.4. Computational complexity

The computational complexity of our proposed cell segmentation method is analyzed in detail as follows. Assuming that the input image has a resolution of $H \times W$, the Gaussian smoothing applied to the image for the purpose of pursuing multi-scale processing has a complexity of $O(H \times W)$. The MHT used to separate foreground cells from the background has a complexity of $O(256)$ since, in this step, the Gaussian smoothing is executed on the image histogram which has a constant dimension, i.e., 256. The DEFA adopted to identify overlapping cells has $O(c^2 \times N)$-level complexity [31], where c is the number of cells and N is the number of foreground cell pixels, respectively. Finally, the multi-scale fusion of cell contours has $O(P)$-level complexity, where P is the total number of pixels

comprising the cell contours. Therefore, in total, the proposed method has a computational complexity of $O(H \times W + c^2 \times N + P + 256)$.

Our cell segmentation method was coded in MATLAB and implemented with an Intel Core i5 1.25 GHz processor. Table 3 details the computational complexity of each step in terms of the processed data size and the CPU time averaged over all analyzed images. In total, our method takes about 7.44 s per image, which compares favorably with the times of the OTSU [30] (about 22.6 s per image) and the RFOVE [29] (about 29.4 s per image).

Table 3. Computational complexity of our proposed cell segmentation method in terms of the processed data size and CPU running time measured at each step and averaged over all analyzed images.

Step	Processed Data Size	Running Time (s)
Image smoothing	$H \times W$	0.10
MHT	256	0.61
Overlapping cell identification	$c^2 \times N$	6.73
Multi-scale cell contour fusion	P	0.002

5. Conclusions and future work

Cell segmentation of fluorescent microscopy images can be done by performing histogram thresholding to separate the foreground cells from the background first, through which the subsequent segmentation steps are greatly facilitated. In the literature, this thresholding operation is commonly conducted at a single scale—that is, the thresholding is implemented after applying a fixed amount of Gaussian smoothing to suppress the jaggedness of the histogram curve, where the scale refers to the standard deviation of the used Gaussian filter. However, it is difficult to determine an appropriate scale for achieving satisfactory thresholding results in different circumstances. To address this crucial problem, an MHT technique has been proposed and investigated. In our MHT technique, the image histogram is smoothed at multiple scales first, and then the smoothed histogram curves are averaged to conduct the histogram thresholding process. To overcome the difficulty of inaccurate segmentation of overlapping cells, the MHT is further incorporated into a multi-scale cell segmentation framework, in which the scale refers to the standard deviation of the Gaussian applied to the input image for noise suppression. In detail, our developed multi-scale framework has three branches, each performing cell segmentation at a given scale based on the MHT, followed by a fusion of the obtained cell contours to arrive at the final segmentation result. Extensive experiments conducted on benchmark datasets have shown that our proposed cell segmentation method can outperform the existing methods by a large margin. In particular, compared to the current state-of-the-art performance level, the average JSC and AJI have been increased from 86.3 to 90.95 (about 5.39% performance gain) and 69.4 to 73.55 (about 5.98% performance gain), respectively.

Our developed multi-scale framework for cell segmentation has the potential to be extended to conduct segmentation on other types of medical images, such as organ images and magnetic resonance brain images. On the other hand, our proposed MHT could also be extended to handle multi-level image segmentation tasks (e.g., the segmentation of breast cancer images) in which the input image

contains multiple targets that need to be segmented at their respective levels. For that, Rosin's method needs to be generalized to process the 2D histogram first, before being further multi-scaled to improve its robustness against noise. All of these issues will be investigated in our future work.

Use of AI tools declaration

The authors declare that they have not used artificial intelligence tools in the creation of this article.

Acknowledgments

This work was supported in part by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant 21KJA520007, in part by the Collaborative Innovation Center of Novel Software Technology and Industrialization and in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. A. Gharipour, A. W. C. Liew, Segmentation of cell nuclei in fluorescence microscopy images: An integrated framework using level set segmentation and touching-cell splitting, *Pattern Recognit.*, **58** (2016), 1–11. https://doi.org/10.1016/j.patcog.2016.03.030
2. D. Riccio, N. Brancati, M. Frucci, D. Gragnaniello, A new unsupervised approach for segmenting and counting cells in high-throughput microscopy image sets, *IEEE J. Biomed. Health. Inf.*, **23** (2019), 437–448. https://doi.org/10.1109/JBHI.2018.2817485
3. F. H. D. Araújo, R. R. V. Silva, F. N. S. Medeiros, J. F. R. Neto, H. C. P. Oliveira, Active contours for overlapping cervical cell segmentation, *Int. J. Biomed. Eng. Technol.*, **35** (2021), 70–92. https://doi.org/10.1504/IJBET.2021.112834
4. Z. Wang, Z. Wang, A generic approach for cell segmentation based on Gabor filtering and area-constrained ultimate erosion, *Artif. Intell. Med.*, **107** (2020), 101929. https://doi.org/10.1016/j.artmed.2020.101929
5. K. Hajdowska, S. Student, D. Borys, Graph based method for cell segmentation and detection in live-cell fluorescence microscope imaging, *Artif. Intell. Med.*, **71** (2022), 103071. https://doi.org/10.1016/j.bspc.2021.103071
6. J. L. Mueller, Z. T. Harmany, J. K. Mito, S. A. Kennedy, Y. Kim, L. Dodd, et al., Quantitative segmentation of fluorescence microscopy images of heterogeneous tissue: Application to the detection of residual disease in tumor margins, *PLoS One*, **8** (2013), e66198. https://doi.org/10.1371/journal.pone.0066198
7. M. Zhao, J. An, H. Li, J. Zhang, S. T. Li, X. M. Li, et al., Segmentation and classification of two-channel C. elegans nucleus-labeled fluorescence images, *BMC Bioinf.*, **18** (2017), 1–13. https://doi.org/10.1186/s12859-017-1817-3

8. Z. Wang, Z. Wang, Robust cell segmentation based on gradient detection, Gabor filtering and morphological erosion, *Biomed. Signal Process. Control*, **65** (2021), 102390. https://doi.org/10.1016/j.bspc.2020.102390

9. M. Salvi, U. Morbiducci, F. Amadeo, R. Santoro, F. Angelini, I. Chimenti, et al., Automated segmentation of fluorescence microscopy images for 3D cell detection in human-derived cardiospheres, *Sci. Rep.*, **9** (2019), 6644. https://doi.org/10.1038/s41598-019-43137-2

10. D. Jia, C. Zhang, N. Wu, Z. Guo, H. Ge, Multi-layer segmentation framework for cell nuclei using improved GVF Snake model, Watershed, and ellipse fitting, *Biomed. Signal Process. Control*, **67** (2021), 102516. https://doi.org/10.1016/j.bspc.2021.102516

11. T. Vicar, J. Balvan, J. Jaros, F. Jug, R. Kolar, M. Masarik, et al., Cell segmentation methods for label-free contrast microscopy: Review and comprehensive comparison, *BMC Bioinf.*, **20** (2019), 1–25. https://doi.org/10.1186/s12859-019-2880-8

12. T. M. S. Mulyana, Herlina, Evenly brightening using kurtosis gaussian pattern to simplify image binarization, in *Journal of Physics: Conference Series*, **1397** (2019), 012076. https://doi.org/10.1088/1742-6596/1397/1/012076

13. P. D. Raju, G. Neelima, Image segmentation by using histogram thresholding, *Int. J. Comput. Sci. Eng. Technol.*, **2** (2012), 776–779.

14. S. Ram, J. J. Rodriguez, Size-invariant detection of cell nuclei in microscopy images, *IEEE Trans. Med. Imaging*, **35** (2016), 1753–1764. https://doi.org/10.1109/TMI.2016.2527740

15. A. A. Ewees, M. A. Elaziz, M. A. A. Al-Qaness, H. A. Khalil, S. Kim, Improved artificial bee colony using sine-cosine algorithm for multi-level thresholding image segmentation, *IEEE Access*, **8** (2020), 26304–26315. https://doi.org/10.1109/ACCESS.2020.2971249

16. X. Yang, R. Wang, D. Zhao, F. Yu, A. A. Heidari, Z. Xu, et al., Multi-level threshold segmentation framework for breast cancer images using enhanced differential evolution, *Biomed. Signal Process. Control*, **80** (2023), 104373. https://doi.org/10.1016/j.bspc.2022.104373

17. S. E. A. Raza, L. Cheung, D. Epstein, S. Pelengaris, M. Khan, N. M. Rajpoot, Mimo-net: A multi-input multi-output convolutional neural network for cell segmentation in fluorescence microscopy images, in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, (2017), 337–340. https://doi.org/10.1109/ISBI.2017.7950532

18. H. Wang, P. Cao, J. Wang, O. R. Zaiane, Uctransnet: Rethinking the skip connections in u-net from a channel-wise perspective with transformer, in *Proceedings of the AAAI Conference on Artificial Intelligence*, **36** (2022), 2441–2449. https://doi.org/10.1609/aaai.v36i3.20144

19. H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, et al., Swin-unet: Unet-like pure transformer for medical image segmentation, in *European Conference on Computer Vision*, (2023), 205–218. https://doi.org/10.1007/978-3-031-25066-8_9

20. D. K. Prasad, M. K. H. Leung, C. Quek, Ellifit: An unconstrained, non-iterative, least squares based geometric ellipse fitting method, *Pattern Recognit.*, **46** (2013), 1449–1465. https://doi.org/10.1016/j.patcog.2012.11.007

21. Z. Shen, M. Zhao, X. Jia, Y. Liang, L. Fan, D. M. Yan, Combining convex hull and directed graph for fast and accurate ellipse detection, *Graphical Models*, **116** (2021), 101110. https://doi.org/10.1016/j.gmod.2021.101110

22. S. Zafari, T. Eerola, J. Sampo, H. Kälviäinen, H. Haario, Segmentation of overlapping elliptical objects in silhouette images, *IEEE Trans. Image Process.*, **24** (2015), 5942–5952. https://doi.org/10.1109/TIP.2015.2492828

23. P. K. Das, S. Meher, R. Panda, A. Abraham, An efficient blood-cell segmentation for the detection of hematological disorders, *IEEE Trans. Cybern.*, **52** (2022), 10615–10626. https://doi.org/10.1109/TCYB.2021.3062152

24. Y. Al-Kofahi, W. Lassoued, W. Lee, B. Roysam, Improved automatic detection and segmentation of cell nuclei in histopathology images, *IEEE Trans. Biomed. Eng.*, **57** (2009), 841–852. https://doi.org/10.1109/TBME.2009.2035102

25. P. Rosin, Unimodal thresholding, *Pattern Recognit.*, **34** (2001), 2083–2096. https://doi.org/10.1016/S0031-3203(00)00136-9

26. J. Liu, J. Xie, B. Li, B. Hu, Regularized cubic B-spline collocation method with modified L-curve criterion for impact force identification, *IEEE Access*, **8** (2020), 36337–36349. http://doi.org/10.1109/ACCESS.2020.2973919

27. J. Antoni, J. Idier, S. Bourguignon, A bayesian interpretation of the L-curve, *Inverse Probl.*, **39** (2023), 065016. http://doi.org/10.1088/1361-6420/accdfc

28. W. V. Drongelen, *Signal Processing for Neuroscientists*, 2nd edition, Academic Press, 2018. https://doi.org/10.1016/B978-0-12-370867-0.X5000-1

29. C. Panagiotakis, A. Argyros, Region-based fitting of overlapping ellipses and its application to cells segmentation, *Image Vision Comput.*, **93** (2020), 103810. https://doi.org/10.1016/j.imavis.2019.09.001

30. M. Liao, Y. Q. Zhao, X. H. Li, P. S. Dai, X. W. Xu, J. Zhang, et al., Automatic segmentation for cell images based on bottleneck detection and ellipse fitting, *Neurocomputing*, **173** (2016), 615–622. https://doi.org/10.1016/j.neucom.2015.08.006

31. C. Panagiotakis, A. Argyros, Parameter-free modelling of 2D shapes with ellipses, *Pattern Recognit.*, **53** (2016), 259–275. https://doi.org/10.1016/j.patcog.2015.11.004

32. B. Zhong, K. K. Ma, On the convergence of planar curves under smoothing, *IEEE Trans. Image Process.*, **19** (2010), 2171–2189. https://doi.org/10.1109/TIP.2010.2046807

33. L. P. Coelho, A. Shariff, R. F. Murphy, Nuclear segmentation in microscope cell images: A hand-segmented dataset and comparison of algorithms, in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, (2009), 518–521. https://doi.org/10.1109/ISBI.2009.5193098

34. N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, A. Sethi, A dataset and a technique for generalized nuclear segmentation for computational pathology, *IEEE Trans. Med. Imaging*, **36** (2017), 1550–1560. https://doi.org/10.1109/TMI.2017.2677499

35. J. P. Bergeest, K. Rohr, Efficient globally optimal segmentation of cells in fluorescence microscopy images using level sets and convex energy functionals, *Med. Image Anal.*, **16** (2012), 1436–1444. https://doi.org/10.1016/j.media.2012.05.012

36. Y. T. Chen, A level set method based on the bayesian risk for medical image segmentation, *Pattern Recognit.*, **43** (2010), 3699–3711. https://doi.org/10.1016/j.patcog.2010.05.027

AIMS Press