

Nucleus segmentation of cervical cytology images based on multi-scale fuzzy clustering algorithm

Jinjie Huang^{a,b}, Tao Wang^{a,b,c}, Dequan Zheng^c, and Yongjun He^b

^aKey Laboratory of Advanced Manufacturing and Intelligent Technology, Ministry of Education, Harbin, China; ^bSchool of Computer Science, Harbin University of Science and Technology, Harbin, China; ^cNetwork and Education Technology Center, Harbin University of Commerce, Harbin, China

ABSTRACT

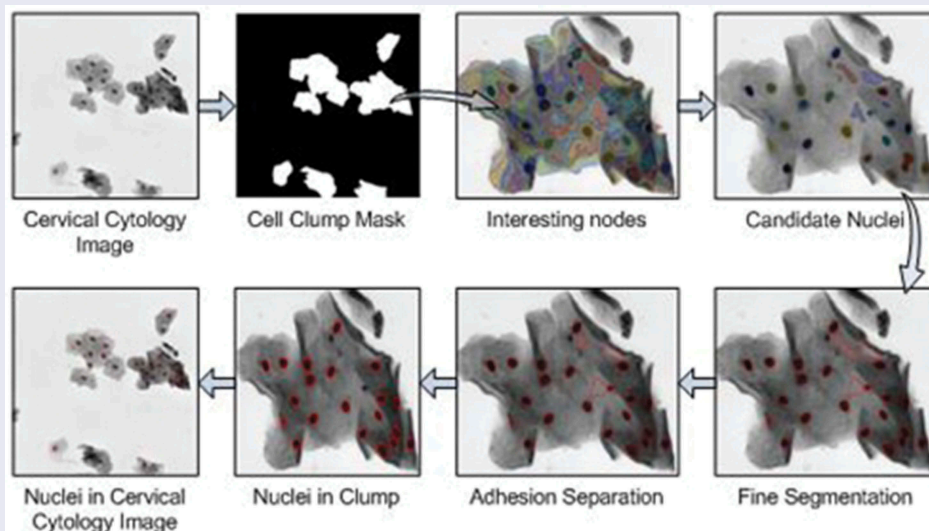
In the screening of cervical cancer cells, accurate identification and segmentation of nucleus in cell images is a key part in the early diagnosis of cervical cancer. Overlapping, uneven staining, poor contrast, and other reasons present challenges to cervical nucleus segmentation. We propose a segmentation method for cervical nuclei based on a multi-scale fuzzy clustering algorithm, which segments cervical cell clump images at different scales. We adopt a novel interesting degree based on area prior to measure the interesting degree of the node. The application of these two methods not only solves the problem of selecting the categories number of the clustering algorithm but also greatly improves the nucleus recognition performance. The method is evaluated by the IBSI2014 and IBSI2015 public datasets. Experiments show that the proposed algorithm has greater advantages than the state-of-the-art cervical nucleus segmentation algorithms and accomplishes high accuracy nucleus segmentation results.

ARTICLE HISTORY

Received 13 January 2020
Revised 23 March 2020
Accepted 24 March 2020

KEYWORDS

Cervical cancer screening; cervical cell; nucleus segmentation; multi-scale fuzzy clustering algorithm




Introduction

Cervical cancer is the second common type of cancer in women, with more than 250,000 women dying from it every year [1]. Fortunately, if early cancerous lesions caused by human papillomavirus HPV are detected in time, cervical cancer can be cured. The cure rate of cervical lesions

depends on the stage of the lesion at the time of diagnosis. If not treated in time, it is likely to cause death.

A pathologist observes a Papanicolaou smear of HPV-positive patient through a microscope, and determines whether there are diseased cells in the image according to the morphology of cervical cells

CONTACT Jinjie Huang ✉ huangjinjie163@163.com  Key Laboratory of Advanced Manufacturing and Intelligent Technology, Ministry of Education, Harbin 150080, China

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

and nuclei. This process is called Papanicolaou (PAP) test. Pap test is one of the most effective detecting methods for cervical cancer through manual screening. However, Pap screening is a time-consuming and repetitive task and requires a high degree of concentration. After a long period of work, the pathologists, who have significant experience in the screening area, may make mistakes due to fatigue and decreased attention. Therefore, a computer-aided diagnosis system (CAD) is needed to assist pathologists in segmenting and identifying cervical cancer cells.

The application of the CAD system reduces the workload of the pathologists and allows them to focus more on the diagnosis and identification of abnormal cervical cells. This improves the accuracy of cervical cancer detection, and further reduces the mortality and incidence of cervical cancer. In response to developing countries with a lack of pathologists, CAD improves the diagnostic efficiency of pathologist; therefore, CAD is also beneficial for the early detection of cervical cancer for developing countries.

When cervical cells are infected, the shape, color, texture, and other characteristics of their nuclei are abnormal. The CAD system must accurately segment the nucleus and determine whether the cells have lesions through feature extraction and classification. The accuracy of nucleus segmentation, as the first step in cancer cell screening, has an effect on the accuracy of the CAD system. A nucleus usually corresponds to a cervical cell; thus, the accuracy of nucleus detection directly affects the subsequent cytoplasm segmentation and recognition. The overlapping, uneven staining, poor contrast, and the presence of neutrophils are also major challenges in cervical nucleus segmentation. Some scholars have studied the nucleus segmentation in single-cell regions and overlapping cervical cell images, respectively.

In the cervical cell images observed under a microscope, the cells are in a state of isolate, adhering or overlapping. Cervical cell samples containing only one isolate cell are shown in [Figure 1](#). Only one boundary (nucleus) or up to two boundaries (nucleus and cytoplasm) need to be detected in these sample images.

Bamford et al. [2] segmented cervical nucleus using active contour method (ACM) based on Viterbi search which could solve the initialization and minimization of the ACM algorithm. Wu et al. [3] utilized a cost function with a hypothetical parameter elliptical shape to detect the nucleus boundary of isolate cells in cervical cytology images. Plissiti et al. [4] applied the local minimum value of the gradient image to identify the position of the nucleus candidate centroids in the Pap smear image, and screened centroids using the color feature of the centroid square neighborhood and the support vector machine classifier. Then, the centroids were regarded as the marker of the watershed algorithm and the false-positive regions were restricted by the binary support vector machine and the shape, texture, and intensity features.

However, the common cervical cells are in adhesion or overlapping cell clumps as shown in [Figure 2](#). We cannot make any assumptions about the number of cells, or expect the cells in the image to be isolated from each other.

In 2014 and 2015, ‘Extended Depth of Field (EDF)-based Overlapping Cervical Cell Segmentation Challenge’ organized by the IEEE International Biomedical Imaging Symposium which published the ISBI2014 and ISBI2015 datasets of overlapping cervical cells, respectively. Many scholars have proposed different methods for nucleus segmentation based on the ISBI2014 and ISBI2015 public datasets. The algorithms for nucleus segmentation are mainly divided into simple linear iterative clustering (SLIC) method [5–8], region-based segmentation method

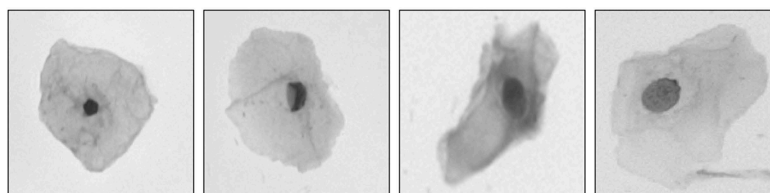


Figure 1. Cervical cell samples containing only one isolate cell.

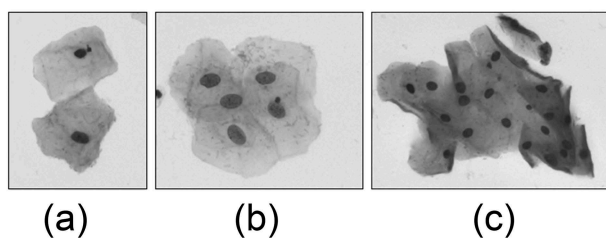


Figure 2. Overlapping or adhering cervical cell Images. Image (a) is an adhering cell clump and image (b), image (c) are overlapping cell clumps.

[7,9–13], convolutional neural network (CNN) [14–17], and clustering method [18–28]. SLIC superpixel algorithm is one of the most popular nucleus segmentation methods currently. It identifies nuclei by generating regular compact superpixels. For the ISBI2014 dataset, Lee et al. applied a triangular threshold method to extract cell clumps and candidate nuclei from superpixels obtained by SLIC method [5]. Afaf et al. [6] constructed the boundary feature vector to classify superpixels. It can solve the problem of over-segmentation of superpixel methods, but the problem of under-segmentation was not solved. For the ISBI2015 dataset, Lu et al. [7] obtained a superpixel boundary map of cervical cytology images by the SLIC method and edge detection algorithm. Afterward, the boundary map was utilized to extract cell clumps and candidate nuclei by the Gaussian mixture model and maximally stable extremal region (MSER) algorithm. Tareef et al. [8] applied a triangular transformation algorithm to identify cell clumps, and combined the SLIC algorithm with a marker-based watershed algorithm to extract candidate nuclei. The SLIC algorithm can accurately segment consistent size targets. The nucleus area at different growth stages has obvious differences. The area of the diseased nucleus is larger than that of the normal nucleus, and can even be six times the size of the normal nucleus. Therefore, it is difficult to accurately extract the nucleus in the cervical cytology image using the SLIC algorithm.

In recent years, region-based segmentation method and CNN have been applied in the field of nucleus segmentation. Jung et al. [9] utilized the MSER algorithm to identify nuclei regions of cervical cytology image. Lu et al. [7] used the MSER algorithm to identify cell clumps and candidate nuclei and applied the nucleus ellipticity feature to screen candidate nuclei. However, in the cervical cytology images with overlapping cells, the

region-based method always presents the false positive and false negative nucleus detection. Song et al. [14] used the CNN algorithm to classify each pixel in a cervical cytology sample into three categories: background, cell clump, and nucleus. This method has greatly improved the accuracy of nucleus extraction. However, the number of parameters that need to be adjusted by the algorithm, the amount of data of the training samples, and a large number of calculations are all problems that cannot be ignored.

Fuzzy C-means clustering (FCM) method [18] was widely used in medical image segmentation, where the FCM algorithm and its improved algorithms were mainly utilized for the segmenting brain MR images [19,20]. Saha et al. [22] adopted first the FCM algorithm for overlapping cervical cytology image segmentation, which added a circular function (CSF) to the FCM algorithm to increase the robustness of the FCM algorithm and improved the boundary segmentation performance. They did further work [21] by adding the MSER algorithm to adaptively calculate the spatial shape force threshold. Compared to the [22], it can get better results based on the recall and precision metrics. However, the segmentation results of the method [21] are dependent on the categories number of clustering. Saha et al. [21] obtained the appropriate categories number of clustering through training set statistics. The selection of the categories number of clustering is a key issue in the performance of FCM and its improved algorithms.

In our paper, we propose a nucleus segmentation algorithm based on multi-scale FCM to address the selection of the categories number of clustering algorithms. Then, we adopt a novel interesting degree based on area prior to measure the interesting degree of the segment. The nodes with appropriate size, uniform intensity distribution, and large solidity can be obtained by the interesting degree based on area prior. Experiments show that the proposed algorithm has greater advantages than the state-of-the-art cervical nuclei segmentation algorithms.

Method

The nucleus segmentation method is mainly divided into three steps in this paper. The first step is to identify the cell clumps, which is to separate the

foreground and background region from the cervical cytology image. In the second step, the multi-scale segments are acquired by the multi-scale FCM algorithm, and then a hierarchical tree is constructed. Afterward, the interesting node in the hierarchical tree is identified in terms of the proposed interesting degree based on area prior. The third step is to achieve nuclei. The DRLSE [29] method is used to finely segment the nucleus boundary and then concave point detection is applied to separate adherent nuclei to get candidate nuclei. Some non-nucleus regions with different shapes and intensity in the candidate nuclei after fine segmentation and adhesion separation are removed by utilizing the feature threshold method to acquire the nucleus region. The process of nucleus segmentation algorithm is shown in Figure 3.

Cell clump extraction

We utilize the cell clump mask to limit the search range for reducing the computational complexity of nucleus segmentation, and separate the cell clump and background regions by the gray value. The FCM algorithm was proposed by Dunn et al. [30] and improved by Bezdek et al. [18]. Bezdek's FCM method has the advantage of a small amount of calculations, so we use the Bezdek's FCM

algorithm [18] to extract the cell clump region in the cervical cytology image.

We define I as the cervical cytology sample. Denote the gray value of the pixel p_i at position i of the sample I as $gray_i, i \in \{1, 2, \dots, P\}$, where P represents the number of pixels in the sample I . We utilize the Bezdek's FCM algorithm to cluster the P gray values into three categories, corresponding to the nucleus, cytoplasm, and background, respectively. Denote v_k as the cluster centers, $k \in \{1, 2, 3\}$. The fuzzy partition matrix is defined as $U = [u_{ik}]_{K \times P}$, where $u_{ik} (1 \leq i \leq P, 1 \leq k \leq K, K = 3)$ represents the membership degree between gray value of pixels i and the category k , and the membership degree should satisfy (1) and (2).

$$u_{ik} \in [0, 1], 1 \leq i \leq P, 1 \leq k \leq K \quad (1)$$

$$\sum_{k=1}^K u_{ik} = 1, 1 \leq i \leq P \quad (2)$$

The objective function of the FCM algorithm is defined as (3), where d is the Euclidean distance of the gray value $gray_i$ and the cluster center v_k .

$$J(U, V) = \sum_{k=1}^K \sum_{i=1}^P (u_{ik})^2 d^2(gray_i, v_k) \quad (3)$$

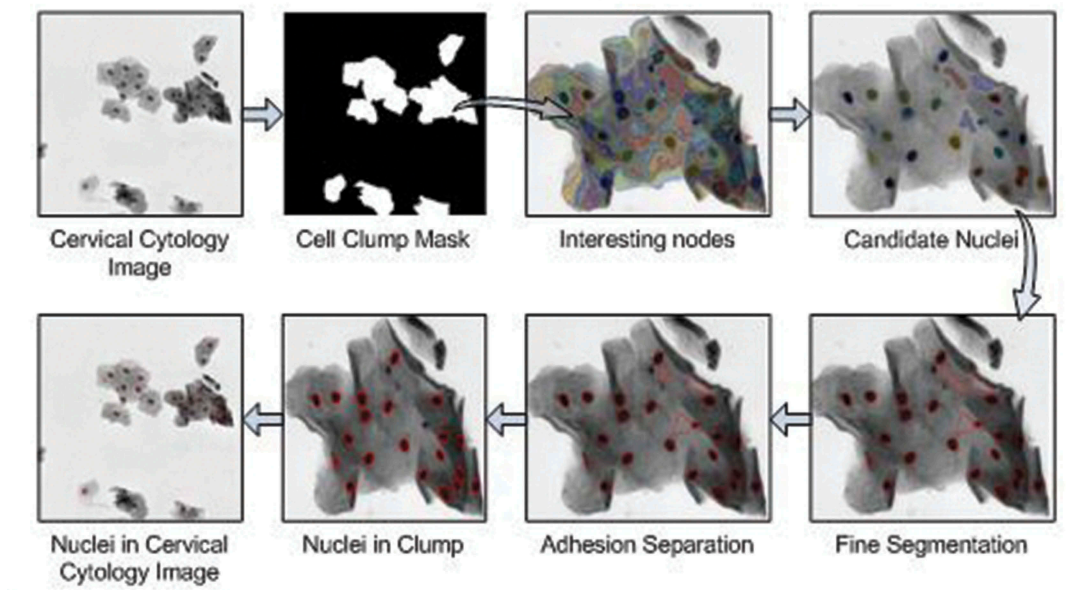


Figure 3. Nucleus segmentation algorithm.

FCM is based on the minimization of the objective function. The cluster centers are updated by (4) iteratively.

$$v_k = \frac{\sum_{i=1}^P (\text{gray}_i u_{ik}^2)}{\sum_{i=1}^P u_{ik}^2} \quad (4)$$

Afterward, the membership degree matrix U is updated by (5).

$$u_{ik} = \frac{1}{\sum_{j=1}^K \left(\frac{\|X_i - v_k\|}{\|X_i - v_j\|} \right)^2} \quad (5)$$

Remarks: In actual cervical cell images, the nucleus, cytoplasm, and background areas generally exhibit three typical levels in grayscale. We take the value of category $K = 3$ in FCM clustering so that the nucleus, cytoplasmic and background regions can be recognized accordingly; On the basis of this, the nucleus region and the cytoplasmic region are combined into a cell clump region; thus, the cell clumps and the background regions can be separated.

Interesting node extraction

Uneven staining, insufficient light, and cell overlap have brought great challenges to the nucleus segmentation of cervical cell clump. Our method utilized three steps to extract the interesting node in the cell clump image. Firstly, the multi-scale FCM algorithm is used to divide the cell clusters at different scales. Afterward, we construct a hierarchical tree based on the inclusion relationship between the segments in the result of the multi-scale division.

Finally, interesting nodes in the hierarchical tree are identified according to the proposed interesting degree based on area prior.

Cell clump segmentation based on multi-scale FCM algorithm

The multi-scale segmentation of the cell clump is to acquire cell clump divisions at different segmentation scales. Different scales can be regarded as segmentation granularity, that is, the smaller the scale, the finer the division of cell clusters. The gray value of the nucleus is generally lower than that of the surrounding cytoplasm, and its intensity of the area is consistent. However, the gray values of cervical cells and nuclei in different locations and growth periods are varied. For example, the surface layer of cervical squamous epithelial cells has a concentrated chromatin in the nucleus region, and its gray value is low, as shown in Figure 4(a); the squamous epithelial cells in the middle layer, by comparison, have loose chromatin and the nucleus gray value is higher than that of the surface nucleus as shown in Figure 4(b); in terms of the overlapping of cells in the cytoplasm area, the cytoplasm may have a lower gray value as shown in Figure 4(c).

The cell clump generally contains multiple cervical cells with different growth periods. In this case, its intensity range is large, and the nucleus region with high intensity may be similar to that of the cytoplasmic region with low intensity caused by overlapping cells. A standard FCM algorithm with clustering category $K = 2$ is used to cluster all the pixels in the cell clump based on the intensity feature, which can only divide the clump into two parts: high-intensity region (cytoplasmic region and lighter nucleus region) and low-intensity

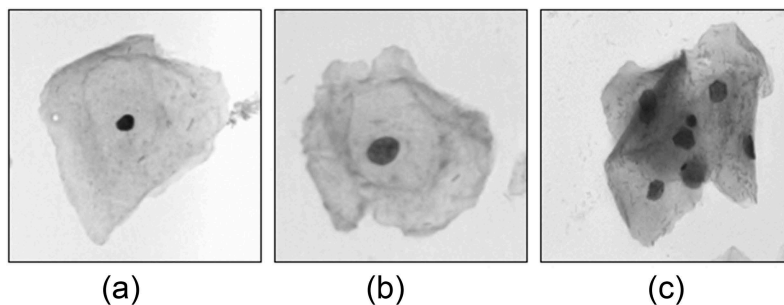


Figure 4. Cervical cells in different locations and growth period. (a) The surface layer of cervical squamous epithelial cells. (b) The squamous epithelial cells in the middle layer. (c) The overlapping of cells in the cytoplasm area.

region (nucleus region and darker cytoplasmic region) as shown in Figure 5. The images in Figure 5(c) are cytoplasm with the low gray value and nucleus with the high gray value, respectively. The FCM algorithm with clustering category $K = 2$ cannot correctly divide the cell clump into nucleus and cytoplasm. By setting the appropriate categories number, the FCM algorithm can better extract the nucleus and cytoplasmic regions. As shown in Figure 6(b), the FCM algorithm with $k = 10$ can segment the nuclei in the cell clump well. The FCM algorithm with clustering category $K = 20$ can also accurately identify the nuclei in the cell clump, but it can be seen from Figure 6(c) that some nuclei are over-segmented.

The selection of categories number is a challenge for FCM algorithm in the nucleus segmentation of cervical cytology images. An appointed categories number applying to the different cells in the cervical cytology image may result in the simultaneous existence of over-segmentation and under-segmentation. We utilize the multi-scale FCM method to avoid the problem of the categories number selection.

If the region R_i^s that is i -th region under the $1/2^s$ scale can be performed $s + 1$ consecutive binary classification from the clump according to divisible principle, it is called divisible region under the $1/2^s$ scale. Under the $1/2^{s-1}$ scale, the divisible regions are divided by FCM with clustering category $K = 2$ and the segmented result is represented as R_a^s , and other region is expressed as R_b^{s-1} . We define the union set $R_a^s \cup R_b^{s-1}$ as the division of cell clump at the scale $1/2^s$, expressed as R^s .

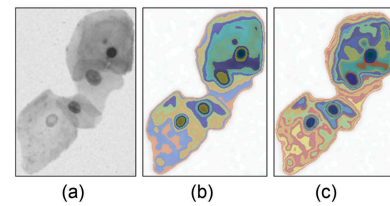


Figure 6. FCM segmentation of the cell clump with clustering category $K = 10$ and $K = 20$. (a) The original image. (b) The segmented result with clustering category $K = 10$. (c) The result with $K = 20$.

The original cell clump can be expressed as R^0 . The divisible principle of this paper is that when the region R_i^{s-1} does not meet the area and intensity thresholds, it can only perform $s-1$ consecutive binary classification. As the scale decreases, the cell clump is divided into smaller segment. When the cell clump division satisfies

$$R^s = \left\{ R_i \mid \sum_{i=1}^N R_i = R, \text{Area}(R_i) \leq \delta_{\text{Area_min}} \text{ and } \text{Gray}(R_i) \leq \delta_{\text{Gray_min}} \right\} \quad (6)$$

under scale $1/2^s$, R^s is the smallest scale division of the cell clump, where N is the number of segments in the cell clump; R is the cell clump area; $\text{Area}(R_i)$ and $\text{Gray}(R_i)$ are the area and average values of intensity of the segment R_i , respectively.

Hierarchical tree construction

To easy analysis attributes of the segments, we construct a hierarchical tree for the segments under different scales. The hierarchical tree can make full use of the inclusion relationship between

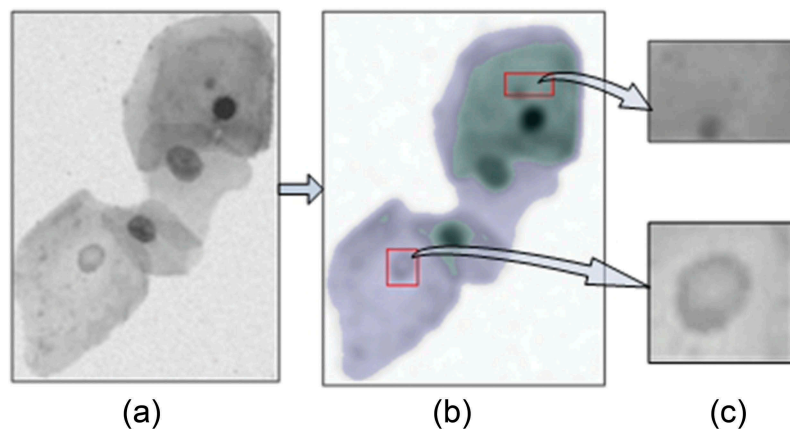


Figure 5. FCM segmentation of the cell clump with clustering category $K = 2$. (a) The original image. (b) The segmented result with clustering category $K = 2$. (c) Incorrect segmentation.

the segments of adjacent scales. The prerequisite for constructing the hierarchical tree is that segments satisfy the nesting structure, which requires that the segment is consistent with or included in another segment in the larger scale. The proposed multi-scale clustering algorithm clusters the segments that meet the divisible principle in the larger scale, though it satisfies the requirements of the nesting structure between adjacent scales.

The segments of each scale are represented as a node of the hierarchical tree. If two nodes have nesting and continuous scale relationship, they are connected by one edge, and the node with large area is expressed as a parent node. Therefore, the root node of the hierarchical tree is the cell clump, and the leaf nodes represent the segments not meeting the divisible principle. Because the cell clump mainly contains the nucleus and cytoplasm, we adopt $K = 2$ as the number of cluster categories. The cell clump is the highest layer of the hierarchical tree, and the leaf nodes are the lowest layer of the tree, as shown in Figure 7.

Recognition of interesting node and selection of candidate nuclei

We identify the interesting nodes from nodes in the hierarchical tree, and then classify them into candidate nucleus and cytoplasmic nodes.

Akçay et al. [31,32] utilized the standard deviation difference and the number of pixels of the nodes to evaluate the interesting degree of nodes, as shown in (7)

$$M(n) = C(n) \times D(n, \text{parent}(n)) \quad (7)$$

where the first term is the number of pixels of the node and the second term is the standard deviation difference between the parent of node parent and itself. The method introduced the number of pixels as another factor to overcome the disadvantages of using only the spectral homogeneity factor. However, these two methods overemphasize the area factor, which tends to identify a node with a large area as an interesting node. According to the area of the nucleus, these two methods are not suitable for measuring the degree of interesting node of the hierarchical tree in this paper.

The nucleus has the features of uniform intensity and large solidity. After the nucleus merges

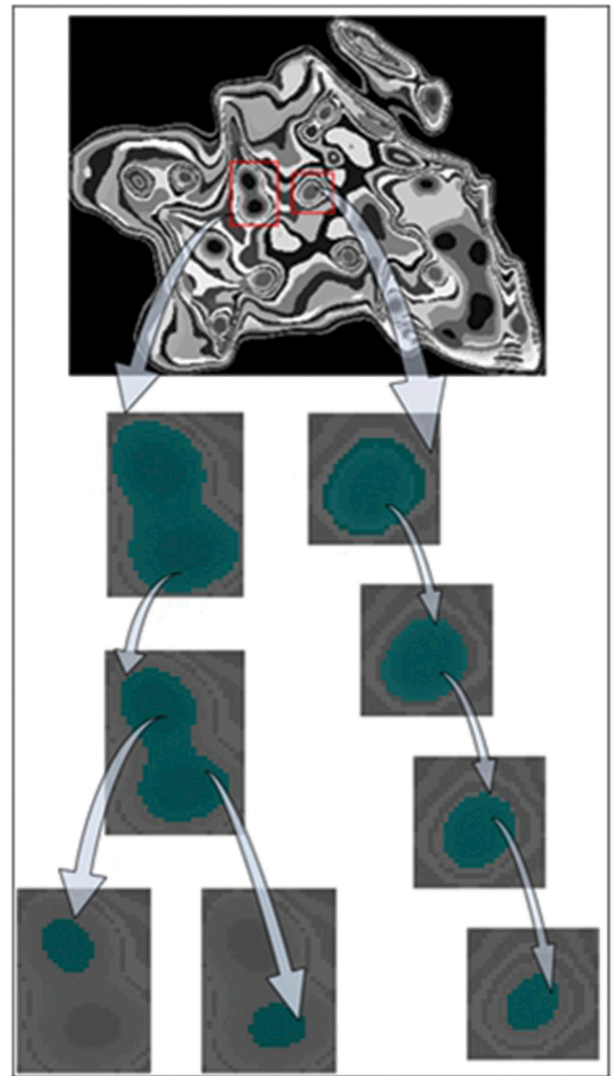


Figure 7. Hierarchical tree structure of cell clusters.

with the surrounding cytoplasm, the intensity value of the region changes drastically. However, in extreme cases, the solidity of a single pixel is the largest. Therefore, we look forward to obtain nodes with appropriate size, uniform intensity distribution, and large solidity.

In this paper, the similarity between nodes is calculated using the intensity consistency measure based on the area prior, which is shown as (8), (9).

$$F(R_1, R_2) = \frac{(s_1 - s_2)^2}{A(R_1) + 1} \quad (8)$$

$$A(R_1) = \begin{cases} 0 & n_1 < \delta_{\text{area_min}} | \delta_{\text{area_max}} < n_1 \\ n_1 & \delta_{\text{area_min}} \leq n_1 \leq \delta_{\text{area_max}} \end{cases} \quad (9)$$

where n_1 represents the number of pixels in R_1 ; and s_1, s_2 represents the variances of R_1 and R_2 , respectively. The solidity S of the region is defined as the ratio of the area of the segment to the area of its convex hull, that is, $S(R) = Area(R)/ConvexArea(R)$. When the shape of the segment is non-concave, its solidity is the largest.

Finally, the interesting degree of node R is defined as shown in (10), and then it is calculated in the hierarchical tree.

$$IOR = F(n, parent(n)) \times S(n) \quad (10)$$

Based on interesting degree, a node is observed when a node with moderate area and high solidity remains the same or merges with similar regions in the higher layer. The following steps are to identify the interesting nodes in the hierarchical tree.

Let N and P be the set of all nodes and their paths in the hierarchical tree, respectively. A subset N^* of N is a set of interesting nodes. Nodes in N^* meet three conditions: 1) the interesting degree of a node must be higher than that of its descendants; 2) any two nodes are not on the same path, that is, any interesting node do not overlap; 3) each path of the hierarchical tree contains only one node in N^* , that is, these interesting nodes constitute a complete cell cluster area.

The interesting node in the hierarchical tree contains the nucleus region expected and cytoplasmic regions. We use support vector machines (SVM) to classify candidate nucleus and cytoplasm regions from interesting nodes. Radial basis function SVM classifier is trained using the features of intensity, solidity, and area in 400 nucleus regions and 2343 cytoplasmic regions from Baseline

dataset [33]. Utilize trained classifier to classify the interesting node to obtain the candidate nucleus, as shown in Figure 8.

Nuclei extraction

The boundary of the candidate nucleus region is around the true nucleus, which cannot accurately describe the true boundary of the nucleus. Therefore, the DRLSE [29] method is used to finely segment the nucleus boundary, and then concave point detection method is applied to separate the adherent nucleus. There are some non-nucleus regions with different features such as shape and intensity in the candidate nuclei after fine segmentation and adhesion separation. These non-nucleus regions are removed by the feature threshold method.

Candidate nuclei refinement

The energy function defined by the DRLSE model is shown in (11), and the external term ε_{ext} of the energy function is defined according to different application scenarios [29].

$$\varepsilon(\phi) = \mu R_p(\phi) + \varepsilon_{ext}(\phi) \quad (11)$$

We utilize a boundary-based active contour model as the external term of the level set evolution algorithm in this paper. Therefore, the energy functional of DRLSE is

$$\varepsilon(\phi) = \mu R_p(\phi) + \lambda L_g(\phi) + \alpha A_g(\phi) \quad (12)$$

where $\phi : \Omega \rightarrow \mathfrak{R}$ is the level set function; $R_p(\phi)$ is a distance regularization term maintaining the sign distance function property of the level set function; $\lambda > 0, \alpha \in \mathfrak{R}$ are the parameters; $L_g(\phi)$ and $A_g(\phi)$

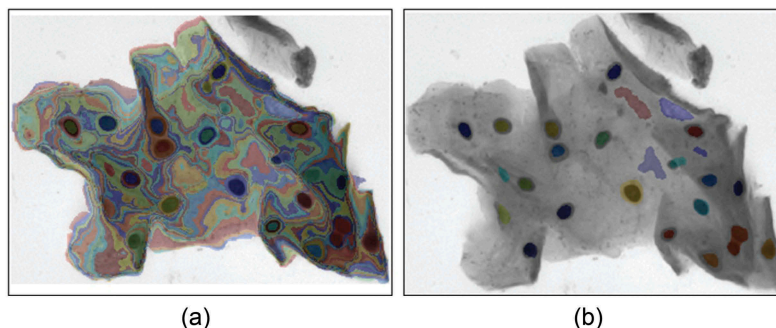


Figure 8. Interesting nodes and candidate nuclei. (a) Interesting nodes. (b) Candidate nuclei.

represent the length functional and area functional, respectively.

Distance regularization term $R_p(\phi)$ is shown in (13), where $p : [0, \infty) \rightarrow \mathfrak{R}$ is defined as a potential function, $p(s) = \frac{1}{2}(s - 2)^2$.

$$R_p(\phi) \triangleq \int_{\Omega} p(|\nabla\phi|) dx \quad (13)$$

Functionals $L_g(\phi)$, $A_g(\phi)$, and $L_g(\phi)$, $A_g(\phi)$, g are defined as follows:

$$L_g(\phi) \triangleq \int_{\Omega} g\delta(\phi)|\nabla\phi| dx \quad (14)$$

$$A_g(\phi) \triangleq \int_{\Omega} gH(-\phi) dx \quad (15)$$

$$g \triangleq \frac{1}{1 + |\nabla G_{\sigma} * I|^2} \quad (16)$$

where δ is the Dirac Delta function; H is Heaviside step function. By Dirac Delta function, the energy $L_g(\phi)$ represents the curve integral of the function g along the zero level set contour of the level set function ϕ . In (16), I represents the image on the domain Ω and g is the boundary indicator function. $|\nabla G_{\sigma} * I|$ represents the gradient value of the image after Gaussian denoising, where G_{σ} is the Gaussian kernel with standard deviation σ and $*$ is the convolution operation for the purpose of noise reduction.

When the contour of the zero level set is at the target boundary, the energy functional gets the minimum value. The area functional $A_g(\phi)$ calculates the weighted area of the region. In special case $g = 1$, the energy function is exactly equal to the area of the region Ω_{ϕ^-} . The length functional $A_g(\phi)$ controls the evolution speed and direction of the zero level set contour. When the initializing contour lies far from the target boundary, the area functional $A_g(\phi)$ has an essential effect. The candidate nucleus is used as the initializing contour of level set in our paper.

In practice, the Dirac Delta function δ and Heaviside step function H are approximated by δ_{ε} and H_{ε} in the L_g and A_g . The δ_{ε} and H_{ε} are defined as (17) and (18).

$$\delta_{\varepsilon} = \begin{cases} \frac{1}{2\varepsilon} [1 + \cos(\frac{\pi x}{\varepsilon})] & |x| \leq \varepsilon \\ 0 & |x| > \varepsilon \end{cases} \quad (17)$$

$$H_{\varepsilon} = \begin{cases} \frac{1}{2} (1 + \frac{x}{\varepsilon} + \frac{1}{\pi} \sin(\frac{\pi x}{\varepsilon})) & |x| \leq \varepsilon \\ 1 & x > \varepsilon \\ 0 & x < -\varepsilon \end{cases} \quad (18)$$

Note that the derivative of H_{ε} is δ_{ε} , that is, $\delta_{\varepsilon} = H'_{\varepsilon}$. The parameter ε is generally set to 1.5. Then, the energy functional is approximated by (19).

$$\varepsilon(\phi) = \mu \int_{\Omega} p(|\nabla\phi|) dx + \lambda \int_{\Omega} g\delta(\phi)|\nabla\phi| dx + \alpha(v) \int_{\Omega} gH(-\phi) dx \quad (19)$$

The minimum value of the energy functional $\varepsilon(\phi)$ is obtained by solving the following gradient flow.

$$\frac{\partial\phi}{\partial t} = \mu \operatorname{div}(d_p(|\nabla\phi|)\nabla\phi) + \lambda \delta_{\varepsilon}(\phi) \operatorname{div}(g \frac{\nabla\phi}{|\nabla\phi|}) + \alpha(v) g \delta_{\varepsilon}(\phi) \quad (20)$$

The candidate nuclei processed by the DRLSE algorithm are shown in Figure 9, where most of them close to the nucleus boundaries and the red boundaries are obtained by the DRLSE algorithm.

Adhesive nucleus segmentation and nucleus screening

The adhesion nucleus segmentation is an important part of the CAD systems. We utilize the concave point detection [34,35] method to segment the adhesive nuclei. Morphological features (geometric center and arc-to-chord ratio) and gradient features (radial symmetry center) are used to determine whether the nuclei are adherent. If the

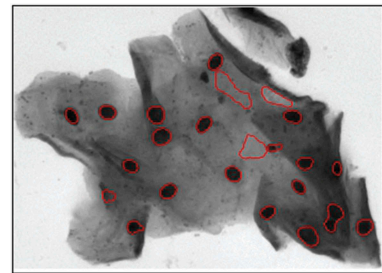


Figure 9. Nucleus boundaries obtained by the DRLSE algorithm. The red boundaries are obtained by DRLSE.

candidate nuclei meet adhesion conditions, the nuclei are called adhesion nuclei. The adhesion conditions are defined as follows:

Condition 1: ($F_s > 1.1$ and $F_r < 4$), Condition 2: ($|r_i - g_i|_2 > 5$)

where r_i is the radial symmetry center; g_i is the geometric center; and $|\cdot|_2$ is the Euclidean distance. The shape parameter F_s is defined as $F_s = L^2/4\pi F_a$, where L is the perimeter and F_a is the area of the candidate nucleus. If the adhesion nuclei satisfy condition 1, the line connecting the two points with the largest arc-to-chord ratio is used as the segmentation boundary to segment the adherent nuclei. If the adherent nuclei satisfy Condition 2, a radial symmetry method [34] is utilized to separate the adherent nuclei. The separation line obtained by the above two methods generally cannot accurately describe the occlusion boundary, which may affect the reliability of nucleus feature extraction. In this paper, these separated regions are used as the initial regions of the DRLSE algorithm, and then the precise nucleus regions are obtained by the level set method post-processing.

There are some pseudo-nuclei in the candidate nuclei, so-called outliers. In this paper, the area, solidity, and eccentricity features are used to screen candidate nuclei. Firstly, the nucleus whose area is smaller than the area threshold is filtered. Then, the candidate nuclei that do not meet the solidity and eccentricity are removed. Finally, the nuclei whose average value of intensity is larger than that of the cell cluster are filtered. The specific threshold value and its acquisition method are given in the experimental section.

Experiment and result analysis

Experimental sample

In this paper, the datasets published by the ISBI2014 and ISBI2015 challenges are used to evaluate the proposed cervical nucleus segmentation method, respectively. The ISBI2014 public dataset contains a total of 945 synthetic cervical cytology images with different cell numbers and overlap rates. The synthetic cell image generation process is as follows: A 40x objective Olympus BX40 microscope and a four-megapixel SPOT Insight camera are used to observe four cervical

smear samples, and 16 non-overlapping fields of view are obtained. Each field contains about 20–60 Pap smear cervical cells. These cells are isolate (do not overlap with other cells) or overlapping with other cells. We can obtain at least 20 focus plane images from each field of view, and the cells show focus or defocus on different focus planes. An overcomplete discrete wavelet transform algorithm is used to generate an EDF (Extended Depth of Field Image) from the multiple focus plane images in a field of view. In the EDF image, every cell in this field of view is in focus. The background of the 16 EDF images and 53 isolate cells is labeled. Utilizing the labeled background and isolate cells, a synthetic cervical cytology image with a resolution of $512 * 512$ is synthesized. The synthesis process is as follows:

Step 1: pixels are randomly selected from the labeled background to constitute the background of the synthetic image;

Step 2: a isolate cell is selected and performed random rigid transformation (rotation, translation scaling) and random linear brightness transformation; the cell passes through the alpha channel (sampling range [0.88, 0.99]); it is placed on the background obtained in step 1.

Step 3: another isolate cell is selected and transformed by the operation of step 2. The added cell need to overlap with the existing cells with one of the following overlap rate ranges [0, 0.1], [0.1, 0.2], [0.2, 0.3], [0.3, 0.4], [0.4, 0.5]. This process is repeated until the number of cells in the synthetic image meets the requirements. The formula for calculating the overlap ratio is

$$overlap_ratio = \max\left(\frac{|A \cap B|}{|A|}, \frac{|A \cap B|}{|B|}\right) \quad (21)$$

where A and B represent two cell regions; and $|\cdot|$ represents the area of the region. In the synthetic process, the nucleus and cytoplasm boundaries are recorded as the artificial labeling results (gold standard) of these 945 synthetic images.

There are 17 samples in the ISBI2015 public dataset, and each sample contains 20 different focus plane images and an EDF image in the same field of view. The image in the dataset is a gray level image with a resolution of $1024 * 1024$. Each sample contains about 40 cervical cells with different overlap ratio, contrast, and

texture. Eight samples in the dataset are defined as the training set. The training set published the nucleus and cytoplasm labeled images. The remaining nine samples are used as the testing set and only the cytoplasm labeled images are published. The ISBI2014 dataset samples and the eight training set samples of ISBI2015 data are used to evaluate the performance of our nucleus segmentation algorithm.

Evaluation method

If nucleus A segmentation results meet $\frac{A \cap B}{A} > 0.6$ and $\frac{A \cap B}{B} > 0.6$, it is said that the nucleus A is correctly detected, where B is the artificial labeling of the nucleus (the gold standard) [34]. The nucleus segmentation results are evaluated by precision Pre_{obj} and recall Rec_{obj} evaluation metrics based on the object. Pre_{obj} and Rec_{obj} are defined as follows:

$$Pre_{obj} = \frac{Cd}{D} \quad (22)$$

$$Rec_{obj} = \frac{Cd}{Gt} \quad (23)$$

where Cd indicates the number of detected nuclei correctly; D indicates the number of detected nuclei; Gt indicates the number of nuclei in the labeled image.

Nucleus segmentation accuracy is evaluated using pixel-based precision Pre_{pix} , recall Rec_{pix} , and Dice coefficient DC , which are defined as follows:

$$Pre_{pix} = \frac{Cd_p}{D_p} \quad (24)$$

$$Rec_{pix} = \frac{Cd_p}{Gt_p} \quad (25)$$

$$DC = \frac{2|A \cap B|}{|A| + |B|}, 0 \leq DC \leq 1 \quad (26)$$

where Cd_p indicates the number of correct detections; D_p indicates the number of detections; Gt_p indicates the number of nucleus pixels in the labeled image.

The nucleus segmentation results also can be evaluated by the positive prediction accuracy

(PPV), negative prediction accuracy (NPV), and the harmonic mean F of between PPV and NPV [14], which are defined as follows [35-39]:

$$PPV = TP/(TP + FP) \quad (27)$$

$$NPV = TN/(TN + FN) \quad (28)$$

$$F = 2 * (PPV * NPV)/(PPV + NPV) \quad (29)$$

where TP indicates the number of pixels correctly predicted as a positive class; FP indicates the number of pixels incorrectly predicted as a positive class; TN indicates the number of pixels correctly predicted as a negative class; FN indicates the number of pixels incorrectly predicted as a negative class; positive type indicates that the pixel is inside the nucleus, and negative type indicates that the pixel is outside the nucleus.

Results

In our nucleus segmentation algorithm, we set $iter_inner = 10$, $iter_outer = 5$ for the DLRSE algorithm. To make the distance regularization term weight μ in (10) satisfy the Courant-Friedrichs-Lewy condition [29], the distance regularization term weight μ is set to 0.04. We set the time step $\Delta t = 5$ to increase the evolution speed. When the difference of variance between areas within the adjacent two evolutionary zero-level set curves is less than δ_{var} , it is considered that the level set function has no longer changed, and the zero level set contour at this time is the best approximation of the nucleus boundary. We set $\delta_{var} = 0.001$ as the stopping condition for the evolution of the level set function.

According to the energy functional of the DRLSE algorithm, the parameters λ and α of the length energy functional $L_g(\varphi)$ and area energy functional $A_g(\varphi)$ have a great influence on the evolution result of the level set function φ . Therefore, the performance of different combinations of λ and α for our algorithm is evaluated in terms of the precision, recall, and DC metric in the ISBI2014 training set. According to the position relation between the candidate nuclei and the true nuclei in the training set, 94% of the candidate nuclei are located inside the real nuclei. The area functional $A_g(\varphi)$ controls the speed and direction

Table 1. The performance of different combinations of λ and α for our algorithm.

λ	α	Pre _{obj}	Rec _{obj}	Pre _{pix}	Rec _{pix}	DC
4	-6	0.993	0.971	0.962(0.05)	0.904(0.08)	0.936(0.03)
4	-4	0.992	0.963	0.976(0.04)	0.882(0.08)	0.924(0.04)
4	-2	0.962	0.961	0.990(0.03)	0.837(0.08)	0.907(0.05)
6	-6	0.991	0.972	0.962(0.05)	0.912(0.07)	0.923(0.03)
6	-4	0.991	0.962	0.983(0.03)	0.885(0.07)	0.935(0.04)
6	-2	0.971	0.961	0.991(0.02)	0.847(0.08)	0.912(0.04)

of the evolution of the level set function. When the parameter α is greater than 0, it evolves into the region and vice versa. Therefore, -2, -4, and -6 are selected as the candidate value for parameter α . The energy functional requires the parameter λ of the length functional $L_g(\varphi)$ is greater than 0, though we choose 4, 6 as the candidate value for the parameter λ . The performance of different combinations of λ and α for our algorithm is shown in Table 1.

From Table 1, when the parameter α is fixed, the parameter λ change has little effect on the final experimental results. For example, when the parameter α is fixed with -4, the experimental result of $\lambda = 6$ is 1% higher than that of $\lambda = 4$ based on the pixel-based precision and DC metrics. However, when the value of the parameter λ is fixed and the parameter α is -2, -4, or -6, we can obtain different experimental results with the various parameter α . The value of the α parameter is set to -2, a high pixel-based precision of 0.99 can be obtained, but the pixel-based recall is low. Therefore, the value of the parameter α parameter can't be set to -2. The algorithm performance with $\alpha = -4$ is not obvious different from performance with $\alpha = -6$, and the DC metric of $\alpha = -4$ is slightly higher than that of $\alpha = -6$. Therefore, the best results are obtained from 45 synthetic cervical cell images in this paper by setting α and λ as -4 and 6, respectively.

We use 900 synthetic cervical cytology images from ISBI2014 and 8 EDF images from ISBI2015 training set as the testing set to evaluate the proposed nucleus segmentation method. The ISBI2014 dataset is synthesized by isolate nuclei in 16 EDF images from ISBI2014 challenge, and some of the EDF images of ISBI2014 are consistent with the EDF images published by the ISBI2015 challenge. Therefore, non-isolate nuclei from the test dataset of ISBI2015 are used as training nuclei for our study. This ensures

Table 2. Parameter settings.

Parameter	Value	Parameter	Value
δ_{area_min}	50	$iter_{inner}$	10
δ_{area_max}	900	$iter_{outer}$	5
$\delta_{solidity}$	>0.85	u	0.04
$\delta_{eccentricity}$	<0.9	Δt	5
δ_{gray_min}	50	δ_{var}	0.001
λ	6	α	-4

that there is no intersection between the training set and the testing set, so that the proposed nucleus segmentation method can be evaluated more fairly. In nine EDF images, a total of 400 nuclei that meet the conditions are used as the training set in this paper. Adjust the thresholds of the area, shape, and solidity features based on the nucleus feature of the training set. All parameter settings in this paper are shown in Table 2.

Based on the 900 synthetic cervical cytology images, the nucleus detection performance and segmentation accuracy of the proposed method are evaluated. We also compared the method to the circular shape constraint FCM algorithm (CiscFC) [21], the methods from ISBI2014 challenge [7,10,36] and nucleus segmentation method proposed by Hady [37], as shown in Table 3. In terms of nucleus detection performance, the object-based precision and recall of our algorithm are higher than that of the five other methods. And with regard to segmentation accuracy, pixel-based recall is 0.980 which is higher than that of other methods, and a competitive DC value is 0.936.

In this paper, experiments are performed on eight EDF images from the training set of the ISBI2015 public dataset, and the experimental results are compared quantitatively with different nucleus segmentation methods under different evaluation metrics, as shown in Tables 4 and 5.

On the ISBI2015 training dataset, our procedure and Lu's procedure code [7] are run on the same computer and software platform. On the basis of Pre_{obj}, Rec_{obj}, Pre_{pix}, Rec_{pix}, and DC metrics, the performance of the proposed algorithm is higher than that of Lu's algorithm (Table 4).

Based on the EDF images from the testing dataset of ISBI2015, the algorithm in this paper is compared with the methods proposed by Hui [11], Zhang [32], Yousef [12], Song [41], respectively, on the three

Table 3. Comparison with the state-of-the-art method.

Algorithm	Pre_{obj}	Rec_{obj}	Pre_{pix}	Rec_{pix}	DC
Ushizima et al [36,40]	0.959	0.895	0.968(0.055)	0.871(0.069)	0.914(0.039)
Nosrati et al. [10,40]	0.903	0.893	0.901(0.097)	0.916(0.093)	0.900(0.053)
Lu et al [7,40].	0.977	0.883	0.942(0.078)	0.912(0.081)	0.921(0.049)
Ratna [21]	0.968	0.882	0.927(0.095)	0.939(0.090)	0.938(0.040)
Hady [37]	0.961	0.933	-	-	-
Our	0.981	0.939	0.864(0.067)	0.980(0.042)	0.936(0.041)

Table 4. Compare with Lu's algorithm.

Algorithm	Pre_{obj}	Rec_{obj}	Pre_{pix}	Rec_{pix}	DC
Lu [37]	0.806	0.781	0.951(0.061)	0.935(0.051)	0.941(0.037)
Our	0.869	0.847	0.942(0.01)	0.92(0.07)	0.927(0.04)

Table 5. Compare with the methods proposed by other scholars.

Algorithm	PPV	NPV	F
Hui ^[11]	0.64	0.81	0.715
Zhang ^[32]	0.83	0.74	0.782
Yousef ^[12]	0.86	0.89	0.875
Song ^[41]	0.95	0.93	0.94
Our	0.88	0.99	0.93

measurement methods of PPV, NPV, and F (Table 5). From Table 5, the nucleus segmentation method described in this paper shows excellent performance under these three metrics, and its performance is generally higher than the first four methods.

The algorithm runs 5 times on the training set of ISBI2014. The purpose is to identify the impact of random initialization of membership matrix and clustering center in the multi-scale FCM algorithm. Record the results of each algorithm run, and calculate the mean and variance, as shown in Table 6. From the results, the variance of the five experimental results is small based on the precision, recall, and DC metrics. Therefore, our multi-scale FCM nucleus segmentation algorithm is less affected by the random initialization of the membership matrix and clustering centers of the FCM algorithm.

Table 6. Experimental results of the algorithm run 5 times on the same dataset.

Algorithm	Pre_{obj}	Rec_{obj}	DC
Run1	0.991	0.965	0.935
Run2	0.992	0.964	0.925
Run3	0.991	0.965	0.938
Run4	0.994	0.942	0.935
Run5	0.991	0.968	0.936
Mean(standard deviation)	0.992 (0.0013)	0.9608 (0.0106)	0.9338 (0.0051)

Discussion

The cervical nucleus segmentation algorithm based on multi-scale FCM can accurately identify the nuclei in the cervical cytology image. From Table 3, on the ISBI2014 training set, the proposed algorithm achieved the best results compared with the five other nucleus segmentation algorithms mentioned in other literatures based on the object-based metric. The precision and recall of nucleus recognition of our algorithm reach 0.98 and 0.94, respectively. The precision of Lu's algorithm is 0.977, which is the best result among the five other methods and is lower 0.4% than that of our method. In the metric of recall, Hady's method of the five other methods obtains the best performance, and recall is 0.933, which is lower 0.63% than that of our method. From Table 4, the algorithm in this paper has achieved good results in terms of object-based precision and recall in the ISBI2015 training set. By comparison with Lu's method, the precision and recall of the object-based algorithm in this paper are 7.2% and 7.8% higher than those of Lu's algorithm, respectively.

The statistical histograms of the experimental results on the ISBI2014 testing dataset and ISBI2015 training dataset under the pixel-based metrics are shown in Figure 10. From Figure 10 (a), in the ISBI2014 testing dataset, pixel-based precision, recall, and DC values above 0.8 account for 0.870, 0.985, and 0.971 of all correctly identified nuclei, respectively. All correctly detected nuclei meet DC value is greater than 0.74, and the proportion of DC values that is higher than 0.95 reaches 34.7%.

According to the above analysis, our algorithm can not only get a high detection rate of nucleus compared to other algorithms but also accurately segment the cell boundary. Figure 11 shows the

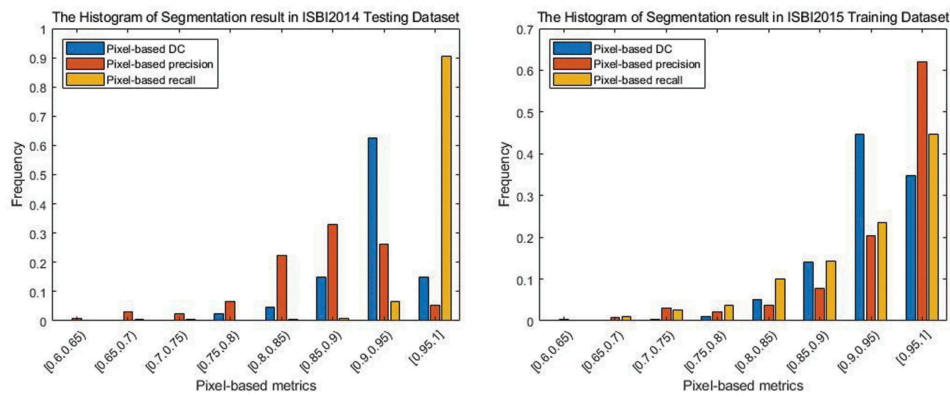


Figure 10. Statistical histograms of pixel-based metrics.

final nucleus segmentation results of our method for synthetic datasets and EDF images.

The detailed results of the nucleus recognition with different numbers of cells and overlapping rate are shown in Table 7 in the ISBI2014 testing dataset. When the number of cells is less than 3 and the overlap rate between cells is less than 0.3, the precision based on the object is 1, which means that the false recognition rate is 0. When the overlap rate is less than 0.2, even if the number of cells reaches 10, the proposed method can still get a lower false recognition result. When the number of cell nuclei is greater than 8 and the overlap rate is greater than 0.4, the error recognition rate of the algorithm reaches a maximum value of 0.08. If the overlap ratio is less than 0.2 among the cells, the recall obtained by the algorithm in this paper is greater than 0.95. As the degree of overlap among cells in the synthetic image increases, the recall rate decreases. When the overlap rate is greater than 0.4, the recall rate decreases obviously. When the number of overlapping cells increases to 9, the algorithm's missed recognition rate reaches 20%. As the number of overlapping cells increases, the contrast decline and the nucleus occlusion cause misidentification. The nucleus gray value is not obvious and is occlusion, as shown in Figure 12, where the rectangular marked area is the case where the nucleus contrast is reduced or occlusion.

This experiment uses the following computer configuration: System: Windows7; Processor: Intel (R) core i5-3470; Memory: 12 GB; Hard disk: 500 G; experimental platform: Matlab R2018a.

For the ISBI2014 synthetic dataset, the average running time processing a cell is 2.5 S, and the average running time processing an image is 15 S. For real EDF images, the running time processing a cell and an image are 10 S and 360 S, respectively.

Conclusion

Nucleus segmentation of cervical smear images is a prerequisite for the widespread application of CAD systems. For the first time, the multi-scale FCM is used to segment the cervical nuclei. It addresses the problem of selection of the categories number of clustering algorithm and avoids the existence of over-segmentation and under-segmentation. Furthermore, a novel interesting degree based on area prior is proposed to measure the interesting degree of the node which can make full use of area and solidity feature of nuclei. The experimental results show that the proposed nucleus segmentation algorithm for cervical smear images has high segmentation precision and the performance of cervical nucleus detection greater than recent state-of-the-art methods.

Highlights

A multi-scale fuzzy clustering algorithm for nucleus segmentation is proposed.

The algorithm avoids selecting the categories number of the clustering algorithm.

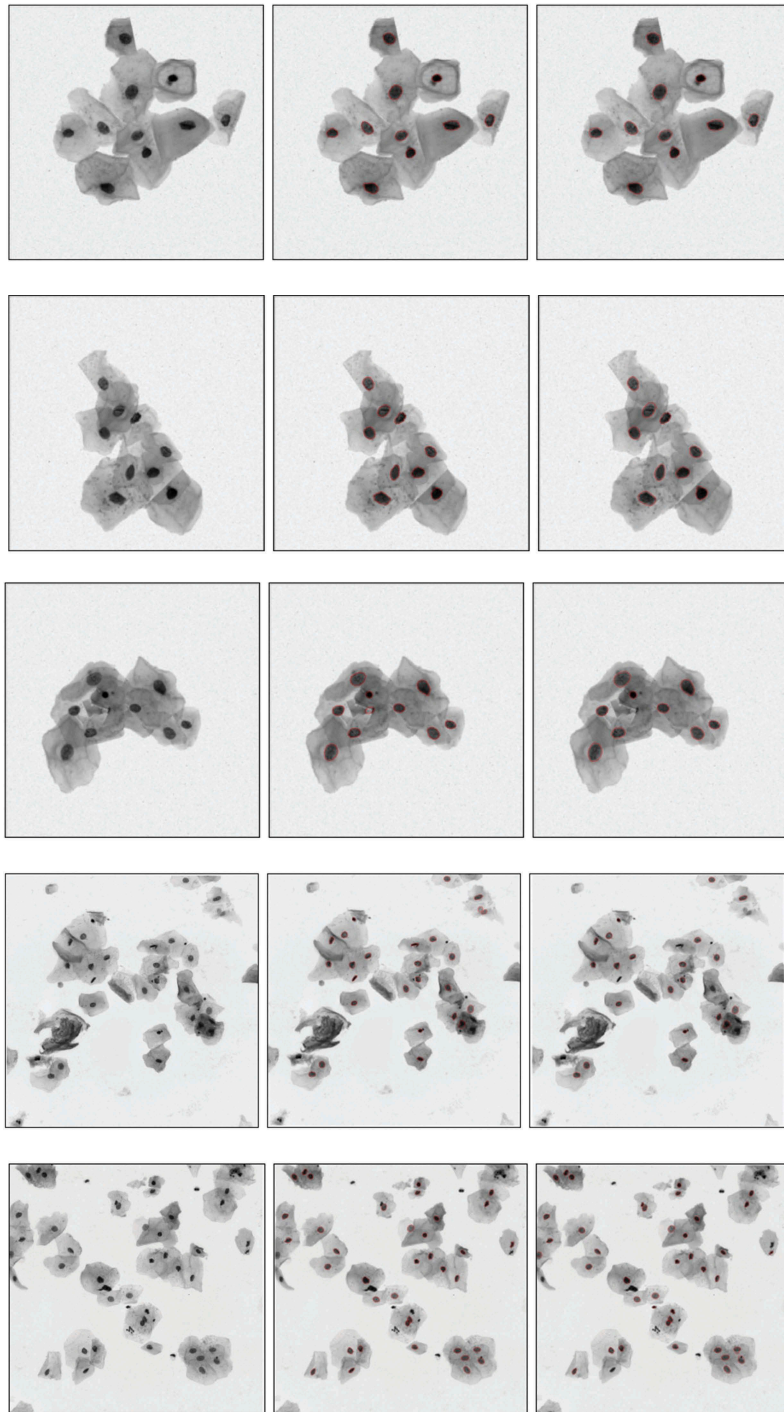


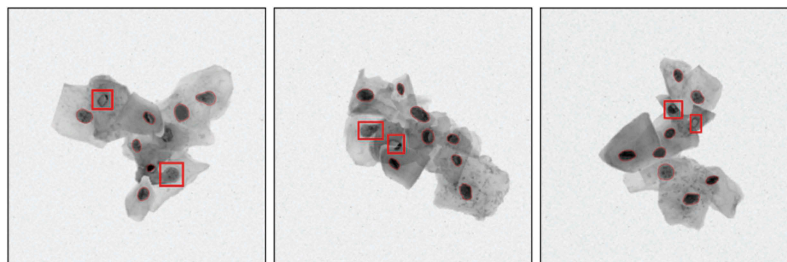
Figure 11. Segmentation results. The left images are the synthetic image and the EDF image, the middle images are the labeled boundary images, and the right images are the nucleus boundary image obtained by the algorithm in this paper.

A novel interesting degree based on area prior is adopted to measure the interesting degree of the segment.

The multi-scale FCM method can address over-segmentation and under-segmentation of FCM algorithm.

Table 7. Experimental results of samples with different cell numbers and overlap rates (object_based precision/object_based recall).

	[0,0.1)	[0.1,0.2)	[0.2,0.3)	[0.3,0.4)	[0.4,0.5)
2 cells	(1)/(0.98)	(1)/(1)	(1)/(0.98)	(0.97)/(0.9)	(0.96)/(0.83)
3 cells	(1)/(0.97)	(1)/(0.98)	(1)/(0.98)	(0.98)/(0.92)	(0.96)/(0.83)
4 cells	(0.99)/(0.95)	(1)/(0.95)	(0.97)/(0.95)	(0.99)/(0.94)	(0.96)/(0.85)
5 cells	(1)/(0.99)	(1)/(0.98)	(0.97)/(0.96)	(0.97)/(0.9)	(0.97)/(0.85)
6 cells	(1)/(0.97)	(0.99)/(1)	(0.98)/(0.91)	(0.98)/(0.91)	(0.95)/(0.82)
7 cells	(1)/(0.99)	(0.99)/(0.99)	(0.99)/(0.97)	(0.98)/(0.87)	(0.94)/(0.88)
8 cells	(0.99)/(0.97)	(0.99)/(0.98)	(1)/(0.93)	(0.99)/(0.91)	(0.95)/(0.87)
9 cells	(1)/(0.97)	(0.99)/(0.98)	(0.98)/(0.97)	(0.96)/(0.88)	(0.92)/(0.8)
10 cells	(1)/(0.99)	(0.99)/(0.99)	(0.97)/(0.93)	(0.99)/(0.90)	(0.92)/(0.83)

**Figure 12.** Contrast reduction and occlusion in the area around the nucleus. The rectangular marked area is the case where the nucleus contrast is reduced or occlusion.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grant 61305001 and the Natural Science Foundation of Heilongjiang Province of China under Grant F201222.

Notes on contributors

Jinjie Huang received his B.Sc. degree in 1990 from Harbin university of Science and Technology, received his M.Sc. degree in 1997 from Harbin university of Science and Technology, and received his Ph.D. degree in 2004 from Harbin Institute of Technology. Now he is a professor in Harbin university of Science and Technology. His main research interests include intelligent control of complex industrial process, pattern recognition and image processing.

Tao Wang received his B.Sc. degree in 2010 from Harbin University of commerce, received his M.Sc. degree in 2013 from Harbin University of commerce. Now he is PhD student in Harbin University of Science and Technology. His main research interests include Image processing and pattern recognition.

Dequan Zheng received his B.Sc. degree in 1991 from Heilongjiang University, received his M.Sc. degree in 1998

from Harbin Institute of Technology, and received his Ph.D. degree in 2006 from Harbin Institute of Technology. Now he is a professor in Harbin university of Commerce. His main research interests include data mining, electronic commerce, and artificial intelligence.

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

References

- [1] World Health Organization, Comprehensive cervical cancer control: a guide to essential practice, World Health Organization, Department of Reproductive Health and Research and Department of Chronic Diseases and Health Promotion, Geneva, Switzerland, 2006.
- [2] Bamford, P, and Lovell B. Unsupervised cell nucleus segmentation with active contours. *Signal Process.* 1998;71.2:203–213.
- [3] Hai-Shan W, Barba J, Gil J. A parametric fitting algorithm for segmentation of cell images. *IEEE Trans Biomed Eng.* 1998;45.3:400–407.

- [4] Plissiti ME, Nikou C, Charchanti A. Automated detection of cell nuclei in Pap smear images using morphological reconstruction and clustering. *IEEE Transactions Inf Technol Biomed.* **2011**;15(2):233–241.
- [5] Lee H, Kim J. Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and cell-wise contour refinement. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops; Las Vegas, USA, 2016.* p. 1367–1373.
- [6] Tareef A, Song Y, Cai W, et al. Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation. *Neurocomputing.* **2017**;221:94–107.
- [7] Lu Z, Carneiro G, Bradley AP. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *IEEE Trans Image Process.* **2015**;24.4:1261–1272.
- [8] Tareef A, Song Y, Huang H, et al. Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. *IEEE Trans Med Imaging.* **2018**;37(9):2044–2059.
- [9] Jung C, Kim C, S W C, et al. Unsupervised segmentation of overlapped nuclei using Bayesian classification. *IEEE Trans Biomed Eng.* **2010**;57(12):2825–2832..
- [10] Nosrati MS, Hamarneh G. A variational approach for overlapping cell segmentation. *ISBI Overlapping Cervical Cytology Image Segmentation Challenge.* **2014**:1–2.
- [11] Jia G, Han G, Jiang J, et al. Dynamic cloud resource management for efficient media applications in mobile computing environments. *Pers Ubiquitous Comput.* **2018**;22(3):561–573.
- [12] Jia G, Han G, Jiang J, et al. Dynamic time-slice scaling for addressing os problems incurred by main memory DVFS in intelligent system. *Mobile Net Appl.* **2015**;20(2):157–168.
- [13] Jia G, Han G, Jiang J, et al. PARS: A scheduling of periodically active rank to optimize power efficiency for main memory. *J Network Comput Appl.* **2015**;58:327–336.
- [14] Song Y, E L T, Jiang X, et al. Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE Trans Med Imaging.* **2017**;36(1):288–300..
- [15] Yin Y, Yueshen X, Wenting X, et al. Collaborative service selection via ensemble learning in mixed mobile network environments. *Entropy.* **2017**;19(7):358.
- [16] Yin Y, Chen L, Yueshen X, et al. location-aware service recommendation with enhanced probabilistic matrix factorization. *IEEE Access.* **2018**;6:62815–62825.
- [17] Yu J, Guo Y, Tao D, et al. Human pose recovery by supervised spectral embedding. *Neurocomputing.* **2015**;166:301–308.
- [18] Bezdek JC. *Pattern recognition with fuzzy objective function algorithms.* Kluwer Academic Publishers, New York, NY: Plenum; 1981
- [19] Ahmed MN, Yamany SM, Mohamed N, et al. A modified fuzzy c-means algorithm for bias field estimation and segmentation of mri data. *IEEE Trans Med Imaging.* **2002**;21(3):193–199.
- [20] Adhikari SK, Sing JK, Basu DK, et al. Conditional spatial fuzzy c-means clustering algorithm for segmentation of {MRI} images. *Appl Soft Comput.* **2015**;34:758–769.
- [21] Saha R, Bajger M, Lee G. Circular shape constrained fuzzy clustering (CiscFC) for nucleus segmentation in Pap smear images. *Comput Biol Med.* **2017**;85:13–23.
- [22] Saha R, Bajger M, Lee G. Spatial shape constrained fuzzy C-Means (FCM) clustering for nucleus segmentation in pap smear images. *2016 international conference on digital image computing: techniques and applications (DICTA); Gold Coast, Australian, 2016.* p 1–8
- [23] Zhang J, Xiao J, Wan J, et al. A parallel strategy for convolutional neural network based on heterogeneous cluster for mobile information system. *Mobile Inf Sys.* **2017**;2017: 1–12.
- [24] Zhang J-L, Yuan J-F, Wan J, et al. Efficient parallel implementation of incompressible pipe flow algorithm based on SIMPLE. *Concurrency Comput-pract Exp.* **2016**;28(6):1751–1766.
- [25] Shrivastava S, Shrivastava L, Bhadauria SS. Performance analysis of wireless mobile ad hoc network with varying transmission power. *Int J Sensor Appl Cont Sys.* SERSC Australia, 2287-8467 (Print); 2207-6344 (Online). May **2015**;3,(1):1–6.
- [26] Yu Y, Rittle LJ, Bhandari V, et al. Supporting concurrent applications in wireless sensor networks. *Proceedings of the 4th international conference on Embedded networked sensor systems.* Boulder Colorado, USA; October, **2006**: 139-152.
- [27] Pradhan S, Sharma K. Cluster head rotation in wireless sensor network: a simplified approach. *Int J Sen App Con Sys.* SERSC Australia, 2287-8467 (Print); 2207-6344 (Online). May **2016**;4(1):1–10.
- [28] Dhakar JS, Parmar M. A survey on grid based data aggregation in wireless sensor networks. *Int J Sen App Cont Syst.* SERSC Australia 2287-8467 (Print); 2207-6344 (Online). May **2016**;4,(1):11–18.
- [29] Li C, Xu C, Gui C, et al. Distance regularized level set evolution and its application to image segmentation. *IEEE Trans Image Process.* **2010**;19(12):3243–3254.
- [30] Dunn JC. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *J Cyber.* **1973**;3: 32–57.
- [31] Akçay HG, Aksoy S. Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Trans Geosci Remote Sens.* **2008**;46(7):2097–2111.
- [32] GençTav A, Aksoy S, Önder S. Unsupervised segmentation and classification of cervical cell images. *Pattern Recogn.* **2012**;45.12:4151–4168.
- [33] Lu Z, Carneiro G, Bradley AP. Automated nucleus and cytoplasm segmentation of overlapping cervical cells. *International Conference on Medical Image Computing and Computer-Assisted Intervention; Springer; Berlin, Heidelberg; 2013.*

- [34] Kong H, Gurcan M, Belkacem-Boussaid K. Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. *IEEE Trans Med Imaging*. 2011;30(9):1661–1677.
- [35] Zhang L, Chen SP, Wang TF, et al. A practical segmentation method for automated screening of cervical cytology. *Proc IEEE Int Conf Intell Comput Bio-Med Inst. Wuhan, China; Dec.* 2011:140–143.
- [36] Ushizima DM, Bianchi AGC, Carneiro CM. *Segmentation of subcellular compartments combining superpixel representation with voronoi diagrams*. No. LBNL-6892E. Berkeley (CA): Lawrence Berkeley National Lab. (LBNL); 2015.
- [37] Phoulady HA, Goldgof D, Hall LO, et al. A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images. *Computerized Med Imaging Graphics*. 2017;59:38–49.
- [38] Li K, Lu Z, Liu W, et al. Cytoplasm and nucleus segmentation in cervical smear images using radiating GVF snake. *Pattern Recogn*. 2012;45(4):1255–1264.
- [39] Zhang L, Kong H, Ting CC, et al. Automation-assisted cervical cancer screening in manual liquid-based cytology with hematoxylin and eosin staining[J]. *Cytometry Part A*. 2014;85(3):214–230.
- [40] Lu Z, Carneiro G, Bradley AP, et al. Evaluation of three algorithms for the segmentation of overlapping cervical cells. *IEEE J Biomed Health Inform*. 2016;21(2):441–450.
- [41] Al-Kofahi Y, Lassoued W, Lee W, et al. Improved automatic detection and segmentation of cell nuclei in histopathology images[J]. *IEEE Trans Biomed Eng*. 2010;57(4):841–852.