Neurocomputing 365 (2019) 157–170

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks☆

Tao Wan[a,b,*], Shusong Xu[a,b], Chen Sang[a,b], Yulan Jin[c], Zengchang Qin[d,*]

[a] School of Biomedical Science and Medical Engineering, Beihang University, Beijing 100083, China
[b] Beijing Advanced Innovation Center for Biomedical Engineering, Beihang University, Beijing 100083, China
[c] Department of Pathology, Beijing Obstetrics and Gynecology Hospital, Capital Medical University, Beijing 100026, China
[d] Intelligent Computing and Machine Learning Lab, School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China

ARTICLE INFO

Article history:
Received 14 January 2019
Revised 24 May 2019
Accepted 23 June 2019
Available online 19 July 2019

Communicated by Pingkun Yan

Keywords:
Nuclei detection
Cytoplasm segmentation
Convolutional neural networks
Overlapping cells
Cervical cytology

ABSTRACT

Accurate cell segmentation is essential for computer-aided diagnosis of cervical precancerous lesions in cytology images. Automated segmentation poses a great challenge due to the presence of fuzzy and overlapping cells, noisy background, and poor cytoplasmic contrast. Deep learning diagnosis technology has showed its advantages in segmenting complex medical images. We present a new framework based on deep convolutional neural networks (DCNNs) to automatically segment overlapping cells in digital cytology. A double-window based cellular detection method is derived to correctly localize individual cells, in which TernausNet is adopted to classify the image pixels into nucleus, cytoplasm, or background. A modified DeepLab V2 model is applied to perform cytoplasm segmentation. To provide more training samples, a synthesis method is utilized to generate cell masses containing touching or overlapping cells. The presented method was tested on three independent data cohorts, including two public datasets. We achieved improved performance in terms of dice coefficient (DSC), false negative and false positive rates, with up to 15% improvement in DSC, compared with the state-of-the-art approaches. The results indicated that the DCNN based segmentation method could be useful in an image-based computerized analysis system for early detection of cervical cancer.

1. Introduction

Cervical cancer is one of the most common causes of cancer mortality in women worldwide [1]. Early detection of precancerous lesions is essential for successful treatment of cervical cancer [2]. Cytology test, such as pap smear test or Thinprep cytologic test, is an important routine screening in diagnosing cervical abnormalities. The diagnostic procedure requires cellular level examination under a microscope by a cytologist or pathologist for searching and detecting nuclear and cytoplasmic atypia, which is tedious and suffers high intra- and inter-observer variations [3]. A computer-assisted diagnosis (CAD) system becomes a solution to automatically and efficiently analyze the digitalized cervical cytological images. Similar to the manual diagnosis, such a system usually considers cell-related image attributes reflecting pleomorphism and chromatism. Therefore, accurate cell segmentation is a vital prerequisite for building a CAD framework. Fully automatically segmenting cervical cells from overlapped clumps is a highly challenging task in computerized cervix cytological analysis.

Recently, many approaches for cervical cell segmentation have been proposed in the literature [4–7]. Most of the current automated methods mainly aimed to segment overlapping nuclei and cellular masses [8–11]. These approaches can be broadly classified into two categories. Majority methods utilized traditional techniques often used in the field of medical image processing or computer vision to perform nuclei and cytoplasm segmentation [12]. For instance, Wang et al. [10] applied the mean-shift clustering and flexible mathematical morphology to split overlapped cell nuclei. Phoulady et al. [13] presented a framework to detect nuclei and segment overlapping cytoplasm in cervical cytology extended depth of field and volume images. The resulting cell boundaries were estimated by a defined similarity metric and then refined via a coarse-to-fine strategy at the pixel level. Nosrati and Hamarneh [14] combined the maximally stable extremal region detector and random decision forest classifier to find the possible nuclei locations. The corresponding cytoplasm was segmented using an active contour model. Yan et al. [15] developed an automatic

☆ This work was partially supported by the Beijing Municipal Natural Science Foundation (Grant no. 7192105), the National Natural Science Foundation of China (Grant nos. 61876197 and 61771325), and Beijing Hospitals Authority YangFan Programme (Grant no. XMLX201820).

* Corresponding author.
E-mail addresses: taowan@buaa.edu.cn (T. Wan), zcqin@buaa.edu.cn (Z. Qin).

https://doi.org/10.1016/j.neucom.2019.06.086
0925-2312/

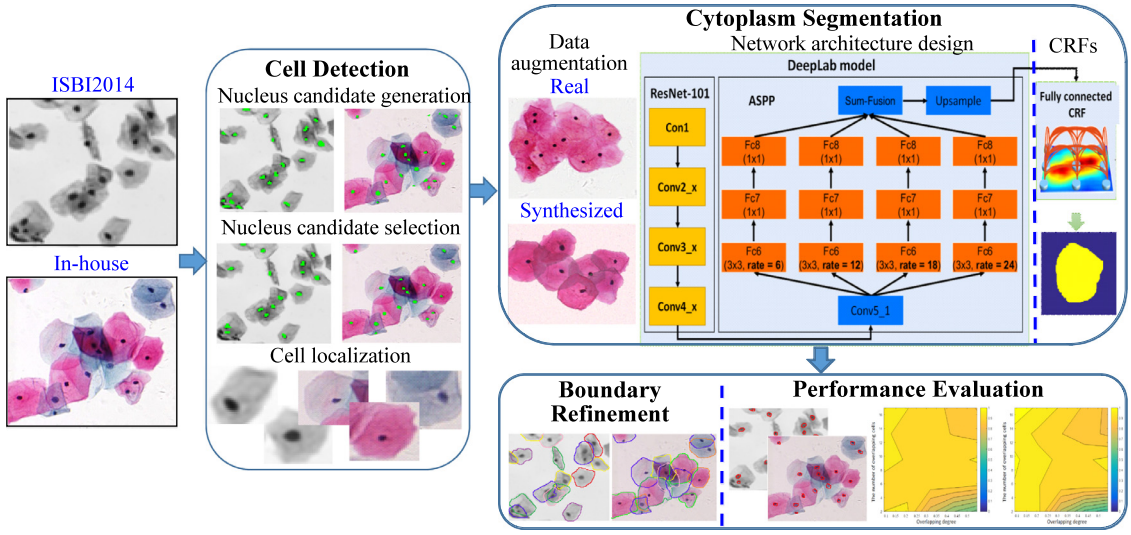


Fig. 1. Workflow of the presented segmentation method based on the DCNN models. Three main modules are designed to perform an automated cell segmentation in cytological images.

cellular image segmentation scheme for RNAi fluorescent genome-wide screening using a multiphase level set method, in which the energy function was formulated based on an interaction model to be able to segment tightly clustered cells with intensity variance and specific phenotypes. Tareefa et al. [16] devised a two-stage segmentation method, in which the segmentation of nuclei and cellular clusters was performed using a set of local discriminative shapes and appearance cues of image superpixels. The cytoplasm was partitioned via a guided sparse shape deformation. In their recent work [12], a multi-pass fast watershed method was adopted to segment both nucleus and cytoplasm from touching or overlapping cell clumps and provided improved segmentation performance. Although all these methods achieved relatively good segmentation results, they made use of prior knowledge of the nucleus-cytoplasm structure at the view of pixel level and could be sensitive to the accuracy of nuclei detection process. The methods may fail in the context of segmenting highly overlapped cytoplasm when the clumped nuclei are two close together. Further, Lu et al. [17] employed multiple level set functions to handle both nuclei and cytoplasm from clumps of overlapping cervical cells. The method obtained improved segmentation performance, but still needed good initializations for the level set functions representing the detection of each cell.

With the recent advent of image processing techniques and advances in computational power, it is now possible to establish deep learning based computer-aided image analysis to facilitate cervical cancer screening and diagnosis [18,19]. In deep learning, a convolutional neural network (CNN) is a class of deep neural network, most commonly applied to analyzing visual imagery [20,21]. For example, Zhu et al. [22] proposed a boundary-weighted domain adaptive CNN for segmenting prostate MR images. A boundary-weighted loss function was designed to make the trained CNN being sensitive to boundaries. Additionally, a boundary knowledge transfer learning was exploited to overcome the challenge caused by the lack of sufficient training data. Song et al. [23] introduced a segmentation framework combining deep learning technique and a deformation model to split cervical cells. A multi-scale CNN was used to learn cervical image pixels and classify each pixel into background, cytoplasm, or nucleus. The detected cell cytoplasm from the CNN was extracted using a multi-cell labeling model. Tareef et al. [24] utilized a deep convolutional neural network (DCNN) model to classify the images into three cellular components (i.e., background, nuclei, and cytoplasmic mass). The cytoplasm contour was delineated via a learning shape prior model. Although the experimental results demonstrated that the deep learning based approaches delivered promising results, these methods encountered the issue of heavy computational burden and generally required a longer processing time when handling whole slide scanned cytological images. This motivated our work to develop a robust and efficient learning based algorithm to automatically segment cervical cells from highly overlapping cell masses in cytological images.

In this paper, we present a new segmentation method for splitting overlapped cells in cervical cytology images based on deep learning models and computerized image analysis techniques. The workflow of the method is depicted in Fig. 1. Similar to the previous work [12,13,16,17], we perform the cellular segmentation via a two-phase framework, in which the regions of interesting identified by the cell detection phase serve as training samples for the subsequent cytoplasm segmentation phase. The algorithm is able to enhance the robustness of the method when processing low quality images with poor contrast and ambiguous foreground/background, such as routinely stained cervical cytology images. To improve the accuracy of cell detection, a three-step approach is utilized to ensure the generation of a set of corrected cellular regions. A TernausNet model [25] is applied to roughly classify the image pixels into nucleus, cytoplasm, or background. The TernausNet model is a modification of the U-Net architecture that is widely used for image segmentation [25]. The model adopts fine-tuning to initialize weights for an encoder of the network. This pre-trained network substantially reduces training time and also helps to prevent over-fitting [25]. This model is particularly useful in medical image segmentation since there are limited amount of labelled data available as training samples. A nucleus candidate selection method is then used to further reduce false positive nuclear pixels. A double-window based cell localization approach is finally employed to determine the overlapped cell regions. Different with the prior DCNN based methods [23,24], in which the cytoplasm segmentation was performed via conventional image processing methods, we conducted a DCNN based instance segmentation to split overlapping cytoplasm in conjunction with cell detection. Compared with the sematic segmentation that does not differentiate each object, the instance segmentation requires the correct detection of individual objects and assignment of different labels to each object instance. Therefore, we are able to identify and extract each cervical cell from touching or overlapping cell masses. A

popular DeepLab V2 model [26] is employed in this work to segment overlapped cells. The DeepLab models have been extensively used in the task of semantic image segmentation and tested on large volumes of image datasets [26]. Further, a modified loss function is devised in order to adjust the model to better deal with cytological images. In summary, the main contributions of this work include:

- A segmentation framework effectively combining DCNN and image processing techniques is designed to accurately separate both nuclei and cytoplasm within highly overlapped cell clumps, which has the potential to be adapted and extended to other types of cytological cells, such as breast fine-needle-aspirates cytology.
- A new loss function is derived by considering cell edge information to force the DCNN model to be more sensitive to the cellular pixels and boundaries, thus leading to improved segmentation performance.

The rest of this paper is organized as follows. Section 2 describes the details of the method. The experimental design and results are presented in Section 3 and Section 4, respectively. The concluding remarks are given in Section 5.

2. Methods

The presented deep learning based segmentation method consisted of three main modules, including cell detection, cytoplasm segmentation, and boundary refinement. A cell detection method was developed to extract the individual cells via the TernausNet model [25] and the double-window based cell localization method. A modified DeepLab V2 model [26] segmented cytoplasm from the image background before performing a post-processing method to refine the cell outer contours.

2.1. Cell detection

2.1.1. Nucleus candidate generation

Nuclei detection is a difficult task due to folded cervical cells, poor contrast of cytoplasm, and presence of neutrophils and artifacts. To reduce the impact of these factors, we first applied a TernausNet based classification method to the entire image. Every pixel was assigned to one of the three labels (nucleus, cytoplasm, and background). This was a preliminary step to roughly find the nucleus candidates and excluded the background and non-nucleus objects. The TernausNet was built based on a classical U-Net architecture and was initialized with weights from a VGG11 encoder pre-trained on a large dataset of ImageNet [25]. Different with the VGG11 model, the TernausNet replaced the fully connected layers with a single convolutional layer of 512 channels serving as a bottleneck central part of the network. The model separated the encoder from the decoder, resulting in advantage of reduced training time and preventing over-fitting. This pre-trained network has showed improved performance in segmenting aerial images compared with the other advanced pre-trained models, such as VGG16 and the network from ResNet family [25]. This model is suitable for the rough nuclei candidate detection on whole-slide scanned cytological images with low cost of computation and reduced overfitting. After classification, we collected all the nucleus labelled pixels as potential nuclei to be used in the subsequent selection process (see Fig. 3(b)).

2.1.2. Nucleus candidate selection

In order to reduce false positive nuclear pixels after the TernausNet based classification, we further ran an Adaboost classifier to filter out the possibly wrong nucleus candidates. Two types of texture features were computed, including the gray level size zone matrix (GLSZM) [27] and histogram of oriented gradients (HoG) [28] on the pixel basis. These image features were used to train the Adaboost classifier.

The GLSZM descriptor is an advanced statistical matrix used for texture characterization based on the fact that homogeneous texture is composed of large flat zones of close gray level. The matrix $\mathcal{M}$ is calculated according to the run length matrix principle. Assume $\mathcal{M}$ has size of $D_g \times D_s$, where D_g is the number of gray levels, and D_s is determined by the size of the largest flat zone. The value of the matrix element m_{ij} is equal to the number of flat zones of size j and of gray level i. The more homogeneous the texture (large flat zones with close gray levels), the wider and flatter the matrix. The GLSZM provides a statistical representation by the bivariate conditional probability density function estimation of the image distribution, meanwhile is robust to the image noise [27]. A total of 13 image attributes were computed based on the GLSZM matric. They are small/large zone emphasis, gray-level/zone-size non-uniformity, zone percentage, low/high gray level, small zone low/high gray level emphasis, large zone low/high gray-level emphasis, gray-level/zone-size variance. The HoG is a popular feature to capture local textural attributes of image. In general, a gradient image $\mathcal{G}$ is obtained via convolving a gradient filter applied on both horizontal and vertical directions. The resulting $\mathcal{G}$ is divided into N small cells without overlaying. For each cell c, a histogram h_c of gradient directions or edge orientations is computed. The HoG feature is the concatenation of all these histograms. The detailed implementations can be found in [27] and [28]. These two types of features worked with the Adaboost classification to further distinguish the true cell nuclei from the previous nucleus candidate generation step. An example is shown in Fig. 3(c).

2.1.3. Cell localization

Based on the locations of identified nuclei, we adopted a double-window method to determine the cellular regions as shown in Fig. 3(d). These regions of interest (ROIs) can form a set of training samples to be used in the following cytoplasm segmentation task. The ROI was defined as a rectangular window with a six-element tuple (Ni_r, Ni_c, Dl_{Ni}, Dr_{Ni}, Du_{Ni}, Dd_{Ni}), where $i \in \{1, 2, \ldots, C\}$, and C is the total number of nuclei centered at nucleus Ni. Ni_r and Ni_c represented the coordinates of nuclear center within the image. Dl, Dr, Du, and Dd denoted the distances between nucleus Ni and the four sides of the window. We defined two initial square ROIs to mark the true cell location, shown in Fig. 2(a) as $R1_{N1}$ and $R2_{N1}$. Two rectangular windows representing two initial ROIs were given as (Ni_r, Ni_c, $D1_{Ni}$, $D1_{Ni}$, $D1_{Ni}$, $D1_{Ni}$) and (Ni_r, Ni_c, $D2_{Ni}$, $D2_{Ni}$, $D2_{Ni}$, $D2_{Ni}$), where $D1 > D2$. We designed a criterion to decide the window size of ROI by considering the overlaid cells. An example is given in Fig. 2(a) showing that two nuclei (N1 and N2) are inside a cell mass. N2's R1 window and N1's R2 window have overlapped region. Dl_{N1} equals to the half of the horizontal distance between N1 and N2. Otherwise, Dl_{N1} equaled to $D1_{N1}$. The same rule is also applied to Dd_{N1}. The resulting ROI is shown in Fig. 2(b). This process ensured that we extracted accurate cell regions containing the overlapping zones between densely clumped cells.

2.2. Cytoplasm segmentation

2.2.1. Data augmentation

Deep learning techniques are supervised learning methods, which require a large amount of training data. In this work, the manual annotations were difficult to perform in the case of densely distributed cells. A synthesis method [17] was utilized to enlarge the image samples in order to train the DeepLab V2 based segmentation algorithm. Moreover, data synthesis can provide a more precise ground truth annotations for individual cell boundary, avoiding

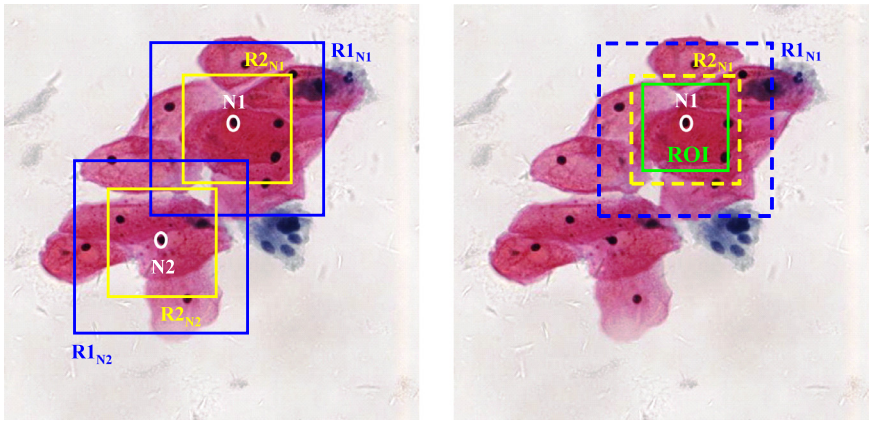


Fig. 2. Illustration of double-window based cell localization method. (a) Two windows ($R1_{N1}$ (blue) and $R2_{N1}$ (yellow)) located at the center of nucleus $N1$; (b) Cell ROI (green) determined by taking into account the positions of $N1$, $N2$, $R1_{N1}$, $R2_{N1}$, $R1_{N2}$, and $R2_{N2}$.

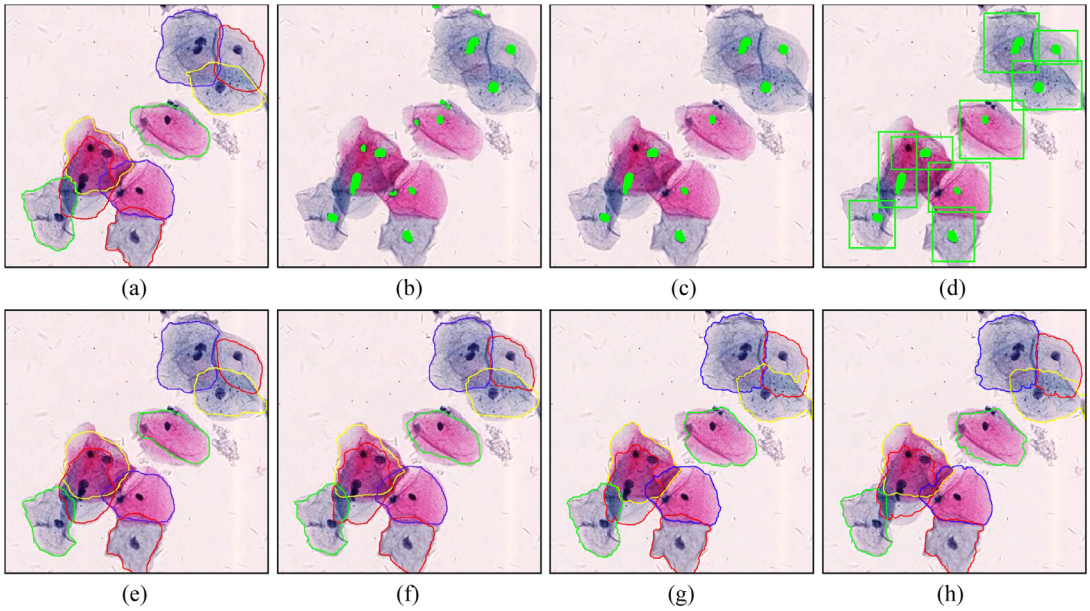


Fig. 3. Image outputs for each step of the presented segmentation method. (a) Original image with manual annotations (ground truth); (b) Nucleus candidate detection (green points); (c) Nucleus candidate selection (green points); (d) Cell localizations indicating by the rectangular windows; (e) Segmentation obtained using the Deeplab V2; (f) Segmentation after CRFs; (g) Final segmentation after the cell boundary refinement using the weighted cross-entropy; (h) Final segmentation using the conventional sum of cross-entropy.

potential error from manual annotations [14]. The method generated the clustered cell images using the single isolated cells. The randomly selected individual cells were used to synthesize various cell masses based on two parameters: 1) the number of cells per image (*NUM*), and 2) overlapping ratio between any pair of cells (*OR*). The detailed synthesis algorithm is described in Algorithm 1.

2.2.2. Network architecture design

We built a popular DeepLab V2 model [26] to segment the cytoplasm from the image background. This model provides a capability in learning multi-scale contextual features through atrous spatial pyramid pooling (ASPP). It allows effectively enlarge the field of view to incorporate larger context without increasing the number of model parameters or the amount of computation. Also, it utilizes an ImageNet-pretrained ResNet-101 [29] as the main feature extractor, which has been proved to be a robust network in the image recognition field. Different from the convolutional ResNet-101, we replaced the model's last block ($Conv5_x$) with the ASPP layer by adding a series of atrous convolutions with different dilation rates, allowing capture multi-scale image attributes. Fig. 1 displays the architecture of the DeepLab V2 model, in which $Conv2_x$, $Conv3_x$, and $Conv4_x$ are repeated [3,4,23] times residual units, respectively. In *Fc*6, these atrous convolutional layers had rates of [6,12,18,24].

The ResNets is a residual learning framework to ease the training of deep convolution neural networks. It contains 4 computational blocks with different number of residual units. These units perform a series of convolutions, which can be defined as

$$y_l = \mathcal{F}(x_l) + x_l, \tag{1}$$

where x_l and y_l are input and output vectors in the lth unit. $\mathcal{F}$ represents the residual mapping that has two or more convolution layers. The formulation can be realized by feedforward neural networks with shortcut connections (skipping one or more layers). The shortcut connections can be simply implemented using an identity mapping, which results in an optimal model having accuracy grains from increased network depth.

Algorithm 1 The cell synthesis algorithm.

Inputs: Original images(O), nuclei masks(NM), cellmasks(CM), NUM, and OR
Output: Synthesis image(S)
Extracting the isolated cells SC_i,($i \in \{1, 2, \ldots, N\}$), and obtaining the background set B_j, ($j \in \{1, 2, \ldots, M\}$) via a random selection of background pixels
Step 1: Placing NUM cells with specified OR on a512×512 canvas
for $i = 1 : NUM$ **do**
  Randomly choosing asingle cell sc_i from SC_i
  if $i = 1$ **then**
    Translatingand rotating sc_i with a random distance and angle
    Randomly generating the other cell positions within a regioncenter on sc_i
  else
    Moving sc_i to (tx, ty) inthe region and then rotating it with a random angle
    Checkingthe validity of the overlapping ratio (or) between sc_i andsc_{i-1} based on OR
    if $or < OR$ **then**
      Adding sc_i onthe canvas
    end if
  end if
end for
Step 2: Computingoverlapping cells' gray values
for $k1 = 1 : NUM$ **do**
  for $k2 = 1 : NUM$ **do**
    Defining a random value $\alpha \in \{0.8, 0.99\}$
    Defining a cell intersecting region Rkk
    $Rkk_{new} = \max(sc_{k1}(Rkk) - sc_{k2}(Rkk), \alpha \times \min(sc_{k1}(Rkk), sc_{k2}(Rkk)))$
  end for
end for
Creating S by adding background from B_j and Gaussian noise

The atrous (or dilated) convolutions allow enlarge the field-of-view of filters at any layer without losing resolution [30]. It introduces a rate f to generate $f-1$ zeros between consecutive filter values. For a $k \times k$ convolution, it effectively enlarges the kernel size of a $k \times k$ filter to $k_{new} = [k + (k-1)(f-1)]$. The ASPP is a module that employs multiple parallel atrous convolutional layers with different rates to learn multi-scale information of image. In this work, the features extracted from each sampling rate were fused to generate a coarse score map showing the rough cell contours.

Moreover, we devised a new loss function to be used in the DeepLab, making the model more suitable to handle cervical cytology images. The original loss function was a sum of cross-entropy between the predicted value and the reference value for every pixel. We found that cell boundaries were often difficult to be accurately detected because of image artifacts and weak gradients along the boundaries. We introduced a weighting factor γ in the loss function in order to force the model to be sensitive to the cell boundaries and cellular pixels enclosed within the cell outer contours. The modified cost function is defined as

$$\mathcal{L} = -\sum_{p=1}^{N} \gamma_p r_p \log(t_p), \tag{2}$$

where t_p and r_p denote predicted and reference values at pixel p, N is the total number of pixels, $\gamma_p = 2$ when p is a boundary or cellular pixel, otherwise $\gamma_p = 1$. Fig. 3(g)(h) displays two examples of cell segmentation using different loss functions. Compared with Fig. 3(h), our modified loss function (shown in Fig. 3(g)) improves the cell segmentation in terms of more precise boundary locations and smoother cellular contours.

2.2.3. Fully connected conditional random fields (CRFs)

The fully connected CRFs [31] was subsequently used as a post-processing method after running the DeepLab model to obtain the final cytoplasm segmentation. The fully connected CRFs combining with deep neural networks have shown improved segmentation performance in both natural and medical images [26,32–34]. As we know, pixel-wise deep learning models classify each pixel independently based on the features derived from the image, producing unsatisfactory segmentation results [35]. The conditional random fields incorporate the prior knowledge of relationships between pixels into the partition process. The fully connected CRFs employ the energy function given as

$$E = \sum_{p} \psi_p(x_p) + \sum_{pq} \psi_{pq}(x_p, x_q), \tag{3}$$

where p and q are the connecting image pixels, x is the label assignment for pixels. The first term can be realized using unary potential $\psi_p(x_p) = -\log(P(x_p))$, where $P(x_p)$ is the label assignment probability at pixel p computed by the DeepLab model. The second term is the pairwise potential having a form that allows for efficient inference while using a fully connected CRF, which can be expressed as

$$\psi_{pq}(x_p, x_q) = \zeta(x_p, x_q) \sum_{m=1}^{2} \omega_m \mathcal{K}_m(p, q), \tag{4}$$

where $\zeta(x_p, x_q) = 1$ when $x_p = x_q$, and zero otherwise, meaning that only nodes with distinct labels are penalized, and ω_m ($m \in \{1, 2\}$) is a predefined weight to balance the two terms. $\mathcal{K}(p, q)$ is formulated using the Gaussian kernel in different feature space:

$$\mathcal{K}_1(p, q) = \exp\left(\frac{\| \mathcal{P}_p - \mathcal{P}_q \|^2}{2\sigma_\alpha^2} - \frac{\| \mathcal{C}_p - \mathcal{C}_q \|^2}{2\sigma_\beta^2}\right),$$
$$\mathcal{K}_2(p, q) = \exp\left(\frac{\| \mathcal{P}_p - \mathcal{P}_q \|^2}{2\sigma_\gamma^2}\right), \tag{5}$$

where $\mathcal{P}_p, \mathcal{P}_q$, $\mathcal{C}_p$, and $\mathcal{C}_q$ are the positions and RGB color for the pixels p, q, respectively. The parameters $\{\sigma_\alpha, \sigma_\beta, \sigma_\gamma\}$ control the scale of the Gaussian kernels. Fig. 3(f) illustrates a segmentation result after performing CRFs. In summary, we combined the ResNet architecture, atrous convolutions, ASPP, and fully connected CRFs into a segmentation network model to achieve superior cytoplasm segmentation with detailed segmentation map along cell boundaries.

2.3. Cell boundary refinement

To obtain precise cell contours, we adopted a distance regularized level set evolution (DRLSE) based segmentation method [36], in which a distance regularization term and an external energy term were derived to drive the motion of the zero level contour toward desired locations. The energy function $\mathcal{E}(\phi)$ is defined as

$$\mathcal{E}(\phi) = \omega_1 \mathcal{R}(\phi) + \omega_2 \mathcal{L}(\phi) + \omega_3 \mathcal{A}(\phi), \tag{6}$$

where ϕ denotes level set function in the image domain, (ω_1, ω_2, ω_3) are the weights to balance the three terms. $\mathcal{R}$ is the regularization term, $\mathcal{L}$ is the distance term, and $\mathcal{A}$ is the region term. They are given by [37]

$$\mathcal{R}(\phi) \triangleq \int_{\phi} R(\nabla\phi) dx, \tag{7}$$

Table 1
The detailed description for three datasets.

Dataset	Image type	Image size	Cells per image	Training no.	Test no.
ISBI2014	Synthetic	512×512	$2-10$	855	90
ISBI2015	Synthetic/real	224×224	$2-10$	945[a]	210
In-house	Real	512×512	$2-20$	460	120

[a] In the experiments, we used 945 synthetic images from ISBI2014 as a training dataset.

$$R(\phi) = \begin{cases} \frac{1}{(2\pi)^2}(1-\cos(2\pi\phi)) \text{ if } \phi < 1, \\ \frac{1}{2}(\phi-1)^2 \text{ if } \phi \geq 1, \end{cases} \tag{8}$$

$$\mathcal{L}(\phi), \triangleq \int_{\Omega} g\delta(\phi)|\bigtriangledown\phi|dx, \tag{9}$$

$$\mathcal{A}(\phi) \triangleq \int_{\Omega} gH(\phi)dx, \tag{10}$$

where H is the Heaviside function, δ is the derivative of H. $g = 1/(1+|1+\bigtriangledown G_\sigma * I|^2)$, where G_σ is the Gaussian kernel with standard deviation of σ. The segmentation maps obtained from the DeepLab model served as initializations for the DRLSE-based segmentation method. Compared with a single pixel as an initial point, the DRLSE method required a small number of iterations to move the zero level set from the initialized boundary to the desired cell boundary. Additional, we assigned ω_3 a small value to avoid the boundary leakage problem. The final cell boundaries, as shown in Fig. 3(g) were identified and obtained after segmentation and refinement through DCNN and level-set methods.

3. Experimental design

3.1. Data description

The presented segmentation method was evaluated on three independent datasets, including two publicly available datasets provided in the first and second overlapping cervical cell segmentation challenges (ISBI2014 and ISBI2015) [17], and one in-house data cohort. ISBI2014 dataset contained 855 synthetic cervical cytology images in the training set and 90 synthetic images in the test set. ISBI2015 had 4 extended depth of field (EDF) images in the training set and 5 EDF images in the test set. We extracted 210 image patches with size of 224×224 pixels from 4 training EDF images. The in-house dataset contained 14 scanned cervical cytology images from 14 patients who underwent a physical examination between 2016 and 2017 at the Beijing Obstetrics and Gynecology Hospital (BOGH). The study was approved by the Ethics Committee of the BOGH and the Beihang University Faculty of Biomedical Science and Medical Engineering, with the requirement for informed consent waived. All the methods were carried out in accordance with the approved guidelines and regulations. All the cases were anonymized. The sides were digitalized via a whole slide scanner (Aperio Technologies Inc., Vista, USA) at $20\times$ optical magnification with a spatial resolution of 0.2456 μm/pixel. We extracted 580 image patches with size of 512×512 pixels from these images. One expert cytologist with more than 15 years of experience marked the locations of nuclei and delineated the enclosed cell contours in the cytology images. The annotations were performed using a self-developed web-based annotating tool providing online environment of collaborative medical image annotation. All the data were accompanied by nucleus and cytoplasm annotations except for the ISBI2015 test dataset, therefore we did not include these images in the evaluation. Table 1 shows the details of the three test datasets.

3.2. Parameter setting

In the nucleus candidate generation, we initialized the learning rage lr as 0.001 for the TernausNet network. The learning rate will decay when increasing iteration number ir according to $lr = lr \times (1 - ir/IR)^{0.9}$, where IR is the total number of iterations. Here, IR was set as 100 for all the experiments. There were two key parameters of the double windows ($D1$ and $D2$) in the cell localization. We set ranges of [70, 100] and [40, 60] for $D1$ and $D2$, respectively, based on the various cell sizes. A series of experiments were conducted using different combinations of {$D1$, $D2$} on three datasets to automatically tune the parameters. We used $\{D1, D2\} = \{90, 55\}$ for the ISBI2014 and in-house datasets, and $\{D1, D2\} = \{85, 55\}$ for the ISBI2015 dataset. In the DeepLab V2 model, the initial learning rate was assigned as 0.0001 and decayed using $lr = lr \times 0.9$ for every 300 iterations. The parameters in the fully connected CRFs were $\{\sigma_\alpha, \sigma_\beta, \sigma_\gamma, \zeta, \omega_1, \omega_2\} = \{3, 80, 13, 1, 1, 1\}$. The technique of stochastic gradient descent with momentum was utilized to optimize the entire training process for both TernausNet and DeepLab models.

3.3. Evaluation metrics

In order to compare with the segmentation results from ISBI2014 and ISBI2015 challenges, we adopted the original evaluation metrics used in both challenges, including dice similarity coefficient (DSC), false negative rate at object level (FNRo), true positive rate at pixel level (TPRp) and false positive rate at pixel level (FPRp). The DSC can be calculated by: $\text{DSC} = 2|R_a \bigcap R_m|/(|R_a| + |R_m|$, where R_a and R_m are the areas enclosed by the automated segmentation and manual segmentation, respectively. $|\cdot|$ is the cardinality of set. $\text{DSC} = 1$ is an indicative of perfect segmentation. For the cytoplasm segmentation, we computed the average DSC over the good cell segmentation, where a good segmentation was defined as $\text{DSC} > 0.7$ per cell. If DSC was lower than a specific threshold ($\tau = 0.3$), we did not account this cell as true positive for the computation of FNRo and TPRp. The higher values of TPRp and the lower values of FNRo and FPRp indicated better segmentation performance. Besides of TPR, we used positive predictive value (PPV) to evaluate the performance of nucleus candidate detection.

3.4. Reference methods

Our method was compared with the other seven alternative approaches. The choice of reference methods took into account both the usage of datasets and implementation availability. Tareef et al. [12] introduced a multipass fast watershed-based method to segment both nucleus and cytoplasm from cell masses using three watershed passes. The method was tested on both ISBI2014 and ISBI2015 datasets. Ushizima et al. [14] derived an automated segmentation method for subcellular compartments through combining superpixel representation with voronoi diagrams. Nosrati and Hamarneh [14] developed a variational method for overlapping cervical cell segmentation based on the maximally stable extremal region detector. This method achieved the best segmentation results in ISBI2014 challenge. Phoulady et al. [13] designed a framework for the boundary approximation of overlapping cells, utilizing

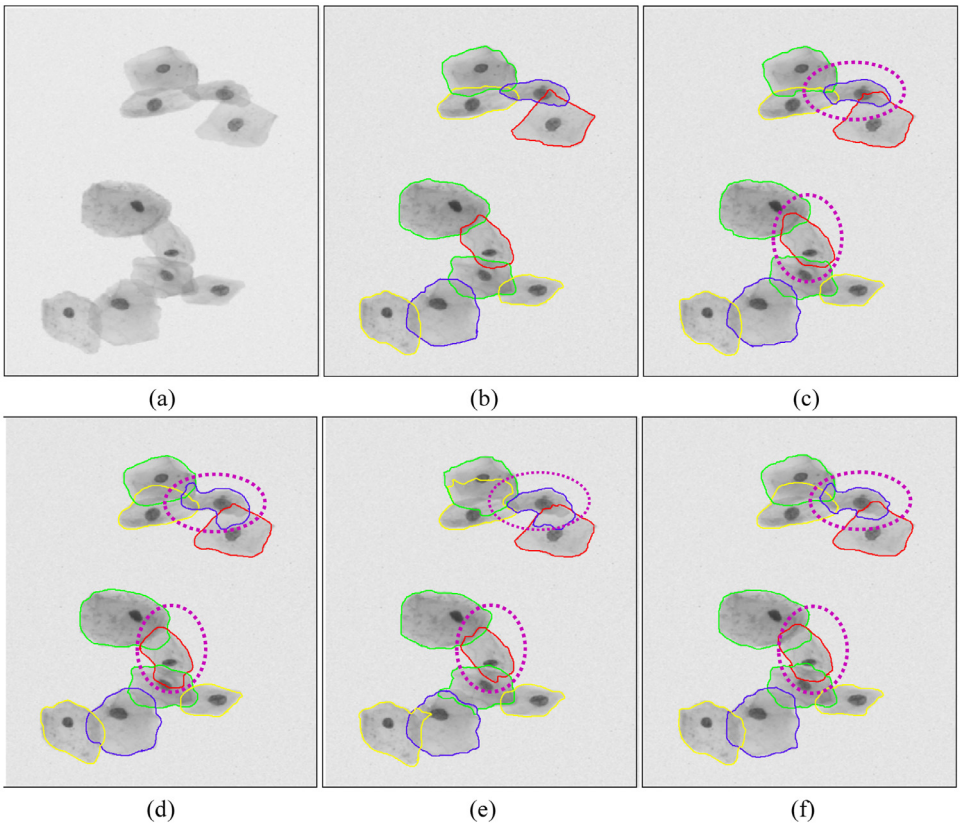


Fig. 4. Segmentation results on ISBI2014 dataset. (a) Original image; (b) Manual annotation; (c) DCNN-based method; (d) Ushizima et al. [14]; (e) Nosrati and Hamarneh [14]. (f) Tareef et al. [12]. The purple dotted circle indicates the differences.

the information in the image stacks to enhance the cell boundary refinements. Their method yielded the top performance in ISBI2015 challenge. Lu et al. [17] utilized a joint optimization of multiple level set functions to separate nuclei and cytoplasm from clumps of overlapping cervical cells. Song et al. [23] employed a multi-scale convolutional neural network to learn cell appearance features, and then applied a deformation model to segment cervical cells. Lee and Kim [38] presented a segmentation method for overlapping cervical cells using superpixel partitioning and cell-wise contour refinement. The original quantitative results reported in the publications were used for comparison.

4. Experimental results and discussions

The DCNN based segmentation method was qualitatively and quantitatively assessed on three independent datasets. The performance was evaluated in comparison with 7 state-of-the-art approaches for segmenting overlapping cervical cells. The effect of parameter variation and computational complexity are analyzed in this section.

4.1. Qualitative evaluation

Figs. 4 –6 show the final segmentation results on the ISBI2014, ISBI22015, and in-house datasets. The reference manual annotations are also displayed along with the automated segmentations. For ISBI2014 dataset, we selected three reference methods (i.e., Ushizima et al. [14], Nosrati and Hamarneh [14], and Tareef et al. [12]) for comparison since these three algorithms achieved the best segmentation performance when attending the ISBI2014 challenge. By carefully examining the figures, we note that the DCNN method (Fig. 4(c)) provides more consistent results with the manual delineations compared with the other three methods. Tareef et al.'s method [12] shown in Fig. 4(f) also yields good segmentation results with the exception of small boundary leakages due to the vague edges within the overlapping cell masses. The rest two reference methods suffered relatively worse boundary leakage problem, resulting in inaccurate cell outer contours, especially on the intersecting areas of overlapping cells.

Fig. 5 illustrates an example of segmentation results applied to the ISBI2015 dataset. Compared with the image shown in Fig. 4(a), this test image (Fig. 5(a)) contains cervical cells with dense distribution and highly overlapping ratio, making the cell segmentation more difficult. We chose three reference approaches (Tareef et al. [12], Phoulady et al. [13], and Lu et al. [17]) for comparison. By observing the figures, we can see that our segmentation demonstrates good capability in separating overlapping cells in complex cervical cytology images due to the accurate nuclei detection via the TernausNet and double-window based cell localization. Tareef et al.'s method [12] performs worse in this case since the method utilizes a watershed-based technique, possibly resulting in over-segmentation when strong and weak image edges appearing at the intersection zones of multiple overlapped cells. In addition, our method is robust and effective in segmenting the clumped cells with inhomogeneous cytoplasm due to the uneven straining. Phoulady et al.'s method [13] yields good segmentation results, but still suffered boundary loss because of the light staining within the cytoplasm. Lu et al.'s method [17] also has the over-segmentation issue on the overlapping areas of clumped cells.

Moreover, the presented method was validated on the in-house dataset. The segmentation results are shown in Fig. 6 along with Nosrati and Hamarneh [14], and Lu et al.'s work [17]. We ran the original implementations from the authors and adjusted their

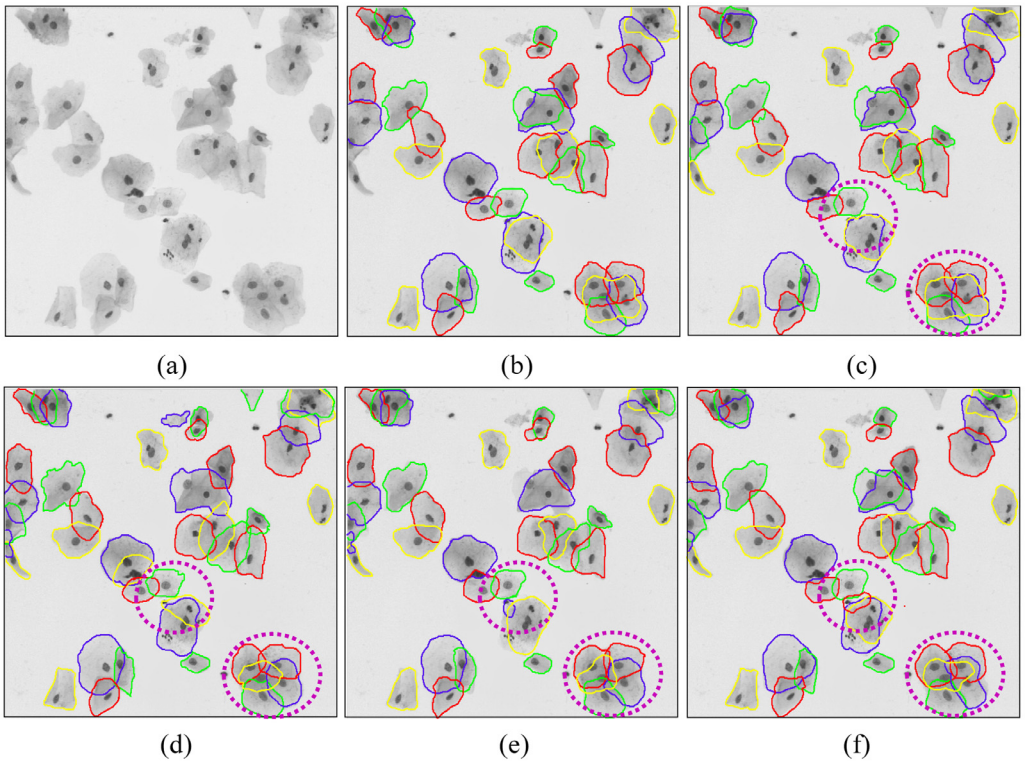


Fig. 5. Segmentation results on ISBI2015 dataset. (a) Original image; (b) Manual annotation; (c) DCNN-based method; (d) Tareef et al. [12]; (e) Phoulady et al. [13]; (f) Lu et al. [17]. The purple dotted circle indicates the differences.

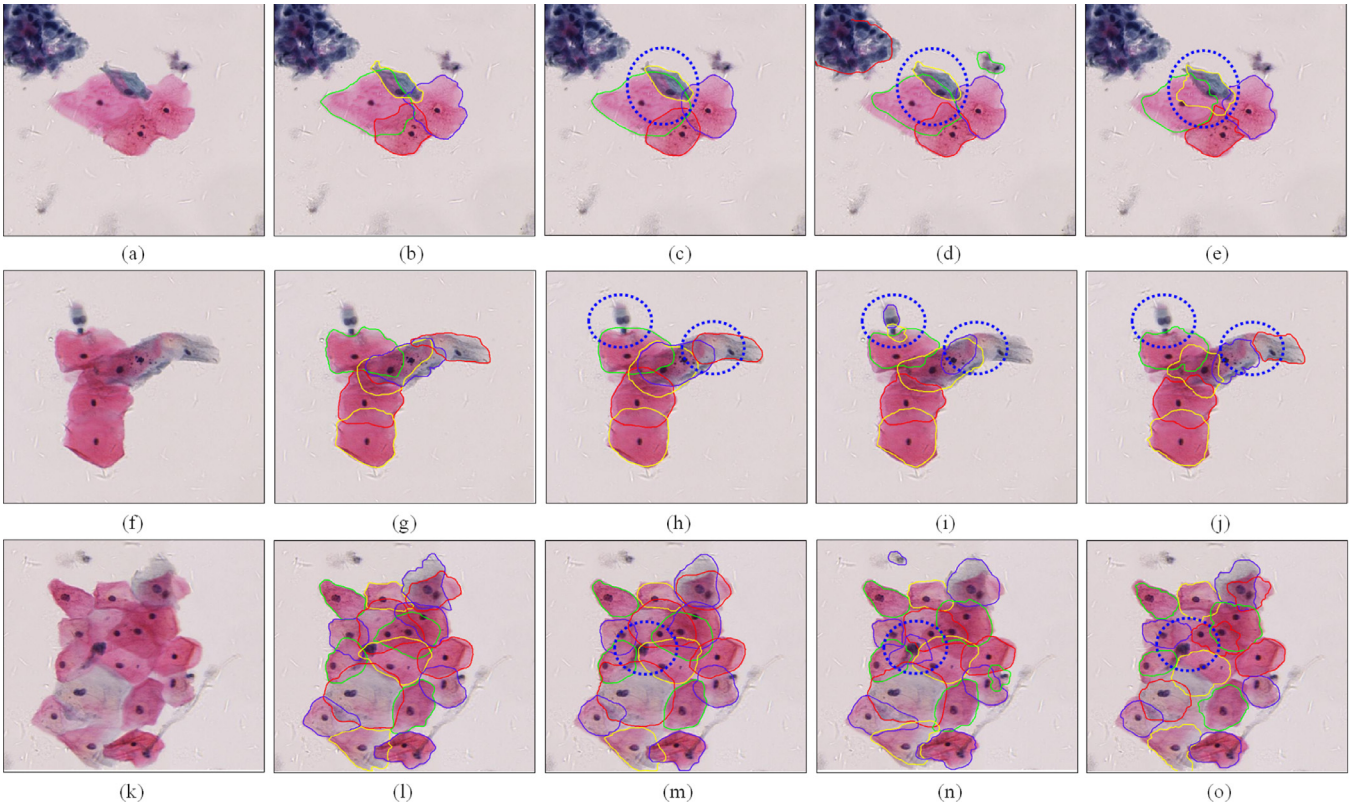


Fig. 6. Segmentation results on in-house dataset. (a)(f)(k) Original images; (b)(g)(l) Manual annotations; (c)(h)(m) DCNN-based method; (d)(i)(n) Nosrati and Hamarneh [14]; (e)(j)(o) Lu et al. [17]. The blue dotted circle indicates the differences.

Table 2
Segmentation performance with different DSC thresholding using DSC, FNRo, TPRp, and FPRp, and the values are in the format of $\mu \pm \sigma$.

Dataset	Thresholding	DSC	FNRo	TPRp	FPRp
ISBI2014	DSC > 0.5	0.94 ± 0.04	0.11 ± 0.18	0.93 ± 0.09	0.001 ± 0.003
	DSC > 0.6	0.94 ± 0.04	0.12 ± 0.15	0.93 ± 0.06	0.001 ± 0.004
	DSC > 0.7	0.94 ± 0.04	0.13 ± 0.11	0.93 ± 0.06	0.001 ± 0.003
	DSC > 0.8	0.94 ± 0.04	0.15 ± 0.17	0.92 ± 0.04	0.002 ± 0.002
ISBI2015	DSC > 0.5	0.92 ± 0.05	0.22 ± 0.14	0.91 ± 0.05	0.001 ± 0.002
	DSC > 0.6	0.92 ± 0.03	0.24 ± 0.21	0.91 ± 0.07	0.001 ± 0.003
	DSC > 0.7	0.92 ± 0.04	0.24 ± 0.19	0.91 ± 0.06	0.001 ± 0.002
	DSC > 0.8	0.91 ± 0.06	0.28 ± 0.24	0.90 ± 0.05	0.001 ± 0.002
In-house	DSC > 0.5	0.92 ± 0.04	0.19 ± 0.11	0.93 ± 0.05	0.003 ± 0.003
	DSC > 0.6	0.92 ± 0.06	0.19 ± 0.09	0.93 ± 0.06	0.003 ± 0.002
	DSC > 0.7	0.92 ± 0.05	0.22 ± 0.14	0.92 ± 0.07	0.003 ± 0.001
	DSC > 0.8	0.91 ± 0.05	0.30 ± 0.10	0.91 ± 0.03	0.004 ± 0.003

parameters to obtain the best performance through the cross-validation strategy. Nosrati and Hamarneh show a precise estimation of the cell boundaries inside the overlapping regions, but less precise in the contours splitting the cells from the background, where inaccurate random forest probability map is the main reason for this issue. Additionally, non-cell regions are partitioned from the background by this method because of the incorrect detection of nuclei. Lu et al.'s approach seems to produce fine segmentation results between clumps and background, but the estimation of the contours in the overlapping cell masses encountered over-segmentation or boundary loss. These problems also happened to the ISBI2015 dataset. Compared with the reference methods, our method achieved superior segmentation in both boundary delineation and cell segmentation within the highly overlapping regions.

4.2. Quantitative evaluation

4.2.1. Nucleus candidate detection

Nuclei detection provided prior knowledge for the cell locations, thus playing a key role in the subsequent cytoplasm segmentation. The final performance was assessed through two measures of TPR and PPV, indicating recall ratio and precision ratio for detecting true nuclei. We used DSC > 0.7 between automated detection and manual delineation to compute true positive and false negative values. Validating our nucleus candidate detection method with three datasets obtained TPR = 0.97 and PPV = 0.99 for the ISBI2014, TPR = 0.96 and PPV = 0.98 for the ISBI2015, and the TPR = 0.98 and PPV = 0.99 for the in-house dataset. The performance evaluation demonstrated that the presented method achieved promising results through a combination of DCNN and conventional image processing techniques.

4.2.2. Cell segmentation

The computation of performance measures, including FNRo, TPRp, and FPRp, depended on the choice of DSC. The DSC here served as a threshold. The lower DSC values led to a larger amount of truly identified cells, thus resulting in higher values of true positive, then better performance in DSC and TPRp. Therefore, we calculated four metrics under different DSC values with a range of {0.5, 0.6, 0.7, 0.8}. The quantitative results are tabulated in Table 2. We can see that as the DSC increases, all the four metrics have stable performance, indicating that our segmentation method is robust at the cell basis. When we set DSC > 0.7, both ISBI2014 and ISBI2015 have almost the same segmentation performance compared to the lower values of DSC threshold. Moreover, the performance drops slightly when increasing the DSC threshold to 0.8 for both datasets. The in-house dataset exhibits larger variations than the other two datasets because the collected images contain a number of cell clumps with denser distribution and higher overlapping ratio.

In addition, we conducted an experiment to evaluate the cell segmentation over different numbers of cells within the cytological images and pairwise overlapping ratios between cells. In this work, the overlapping ratio is defined as $max(|R_A \bigcap R_B|/|R_A|, |R_A \bigcap R_B|/|R_B|)$, with R_A and R_B representing the regions enclosed within the cells A and B, and $|\cdot|$ denoting the area of the region. The overlapping ratio indicates the degree of interaction between a pair of cells inside the cell masses. We analyzed the segmentation performance in terms of DSC and TPRp by varying the cell numbers and the overlapping ratios under DSC > 0.7. Fig. 7 shows a graphical visualization of segmentation results with respect to various number of cells and degrees of overlap. These results show that our methodology can successfully segment cell masses containing a large number of cells (between 10 and 16 cells) even when the overlapping ratio is relatively high (between 0.4 and 0.5). This is due to the fact that the segmentation model is forced to learn the characteristics of the cell boundaries through the improved loss function formulation. We also note that with the increased values of overlapping ratio, the performance drops when the number of cell is small. This is mainly due to the inaccurate detection of nuclei when two cells are too closed to be split. The results suggested that our method was robust and effective to segment cytology images in the context of highly overlapping cell clumps.

4.2.3. Comparison evaluation

The comparison of segmentation performance was quantitatively evaluated using four measures (DSC, FNRo, TPRp, and FPRp). We employed the same DSC threshold, in which a segmented cell in the ground truth was considered to be missed if there was no region in the segmentation result that had a DSC greater than 0.7 with it. The experimental results are listed in Table 3. Nosrati & Hamarneh [14] and Lu et al. [17] used the ISBI2014 dataset. Song et al.'s method [23] and Phoulady et al.'s method [13] were assessed using the ISBI2015 dataset. Tareef et al. [12] and Lee & Kim [38] utilized both ISBI2014 and ISBI2015 datasets. In the table, we adopted all the measure values reported in their originally published work. For the in-house dataset, the original implementations of Nosrati & Hamarneh [14] and Lu et al. [17] were executed to perform the cell segmentation.

By examining the results shown in Table 3, our method achieved the best segmentation performance in terms of DSC and FPRp for both ISBI2014 and ISBI2015 datasets, and highest FNRo and TPRp for ISBI2014 and second best FNRo for ISBI2015. For the in-house dataset, our method outperformed the other two reference methods in all the metrics. The DCNN based method yielded greater DSC measure with up to 6.90% and 6.20% improvements

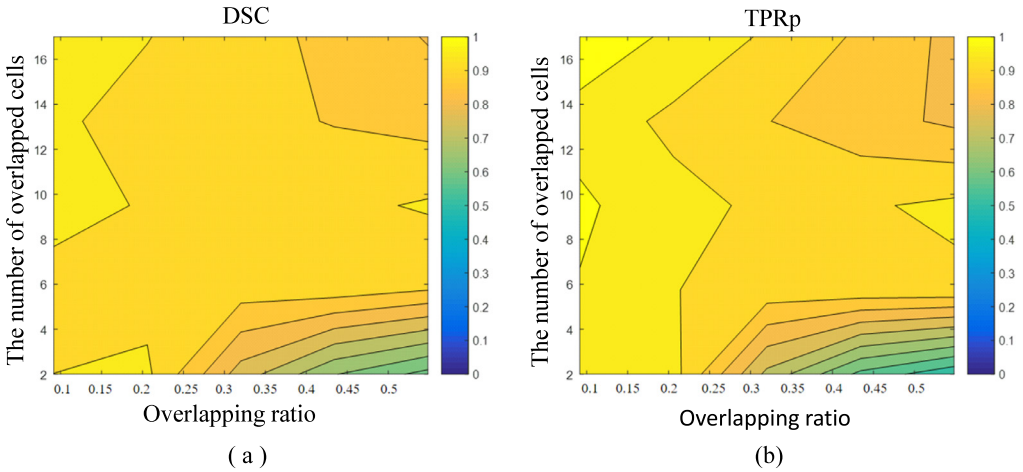


Fig. 7. The graphical visualization of segmentation results using (a) DSC and (b) TPRp with respect to various numbers of cells and overlapping ratios on the in-house dataset. The results suggest that our method is robust and effective to separate highly overlapping cells.

Table 3
Comparison of segmentation performance on three datasets using DSC, FNRo, TPRp, and FPRp, and the values are in the format of $\mu \pm \sigma$. The bold text indicates the best performance.

Method	DSC	FNRo	TPRp	FPRp
ISBI2014				
Tareef et al. [12]	0.89 ± 0.07	0.27 ± 0.28	0.91 ± 0.09	0.004 ± 0.005
Nosrati & Hamarneh [14]	0.87 ± 0.08	0.14 ± 0.17	0.90 ± 0.09	0.005 ± 0.004
Lee & Kim [38]	0.90 ± 0.08	0.14 ± 0.19	0.88 ± 0.10	0.002 ± 0.002
Lu et al. [17]	$0.88 \pm N/A$	$0.21 \pm N/A$	$0.92 \pm N/A$	$0.002 \pm N/A$
The DCNN method	$\mathbf{0.93} \pm 0.04$	$\mathbf{0.11} \pm 0.13$	$\mathbf{0.93} \pm 0.05$	$\mathbf{0.001} \pm 0.002$
ISBI2015				
Tareef et al. [12]	0.85 ± 0.07	$\mathbf{0.11} \pm 0.17$	$\mathbf{0.95} \pm 0.07$	0.004 ± 0.004
Song et al. [23]	$0.89 \pm N/A$	$0.26 \pm N/A$	$0.92 \pm N/A$	$0.002 \pm N/A$
Lee & Kim [38]	0.88 ± 0.09	0.43 ± 0.17	0.88 ± 0.12	0.001 ± 0.001
Phoulady et al. [13]	0.85 ± 0.08	0.16 ± 0.22	0.94 ± 0.06	0.005 ± 0.005
The DCNN method	$\mathbf{0.92} \pm 0.05$	0.13 ± 0.15	0.91 ± 0.05	$\mathbf{0.001} \pm 0.003$
In-house dataset				
Nosrati & Hamarneh [14]	0.78 ± 0.39	0.56 ± 0.28	0.81 ± 0.12	0.071 ± 0.037
Lu et al. [17]	0.84 ± 0.09	0.32 ± 0.24	0.88 ± 0.10	0.032 ± 0.024
The DCNN method	$\mathbf{0.92} \pm 0.04$	$\mathbf{0.15} \pm 0.13$	$\mathbf{0.93} \pm 0.05$	$\mathbf{0.013} \pm 0.012$

compared with the reference methods for ISBI2014 and ISBI2015, respectively. Also on the in-house dataset, our method obtained a comparable DSC (0.92) with the ISBI datasets while the performance of the reference methods decreased from 0.87 to 0.78 for Nosrati & Hamarneh [14], and from 0.88 to 0.84 for Lu et al. [17]. Also, the method achieved great improvement in lowering the FNRo, which denoted the rate of cells missed in the ground truth. Tareef et al.'s approach [12] yielded a high TPRp but along with a relative low DSC, which reflected inaccurate segmentation due to the boundary leakage, causing high true positives as well as high false positives. Our method outperformed Lu's method [17] in both ISBI2014 and in-house dataset, suggesting that the integration of DCNN model and traditional image processing techniques could provide good segmentation for largely clustered and highly overlapping cells. Moreover, an evaluation on three different datasets demonstrated the superiority of the presented algorithm over the other state-of-the-art methods.

4.3. Analysis of influence factors

4.3.1. Effect of parameter variations

In the phase of cell localization, the window size controlled the extraction of cellular regions, and further affected the subsequent cytoplasm segmentation. Therefore, the choice of double window parameters [$D1$, $D2$] was essential for evaluating the final segmentation performance. We conducted a series of experiments using different combinations of [$D1$, $D2$] on three datasets to automatically tune the parameters. Fig. 8 illustrates the three surface plots using the DSC metric by varying $D1$ from 70 to 100, and $D2$ from 40 to 60 on three data cohorts. The search process was ran automatically through 5-fold cross-validation to find the most appropriate combination of parameters for best performance. The values of [$D1$, $D2$] obtaining the top segmentation performance were adopted in the experiments. The figures suggested that the performance of our segmentation approach was resilient to the choice of the parameters of double windows.

4.3.2. Impact of loss function on segmentation

We modified the loss function of the DeepLab by introducing a weighting factor to force the model to assign large weights to the boundary or cellular pixels. To analyze the impact of our devised loss function on the performance of segmentation, we compared the weighted cross-entropy with the sum of cross-entropy (without a weighting factor) and the Dice loss on three test datasets. Four metrics (DSC, FNRo, TPRp, and FPRp) were computed to quantitatively measure the segmentation performance using different loss functions. Similarly, $DSC > 0.7$ was set to calculate true positive and false negative values. From Table 4, it can be observed that using the presented loss function helps to achieve better performance than using the conventional cross-entropy and Dice loss.

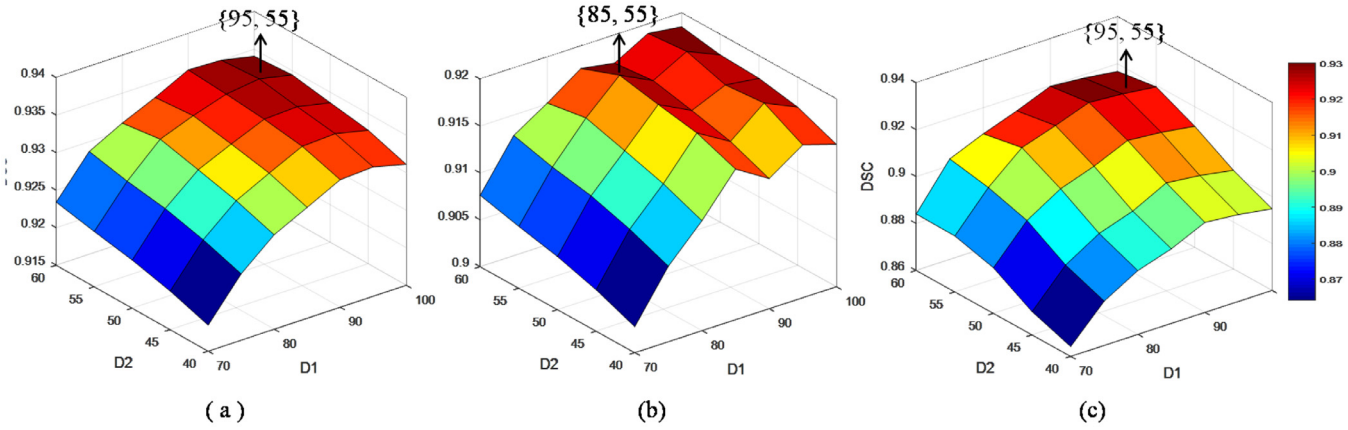


Fig. 8. Surface plots of segmentation performance (dice similarity coefficient (DSC)) using different values of *D*1 and *D*2. (a) ISBI2014; (b) ISBI2015; (c) In-house dataset.

Table 4
Effect of loss functions in segmentation performance on three datasets, and the values are in the format of $\mu \pm \sigma$. The bold text indicates the best performance.

Loss function	DSC	FNRo	TPRp	FPRp
ISBI2014				
Weighted cross-entropy	**0.93** ± 0.04	**0.11** ± 0.13	0.93 ± 0.05	**0.001** ± 0.002
Sum of cross-entropy	0.92 ± 0.05	0.13 ± 0.16	0.93 ± 0.06	0.002 ± 0.002
Dice loss	0.90 ± 0.06	0.19 ± 0.20	0.91 ± 0.07	0.003 ± 0.003
ISBI2015				
Weighted cross-entropy	**0.92** ± 0.05	**0.13** ± 0.15	**0.91** ± 0.05	**0.001** ± 0.003
Sum of cross-entropy	0.91 ± 0.07	0.21 ± 0.11	0.90 ± 0.04	0.003 ± 0.003
Dice loss	0.91 ± 0.07	0.17 ± 0.13	0.89 ± 0.06	0.002 ± 0.003
In-house dataset				
Weighted cross-entropy	0.92 ± 0.04	**0.15** ± 0.13	**0.93** ± 0.05	**0.013** ± 0.012
Sum of cross-entropy	0.92 ± 0.06	0.19 ± 0.23	0.92 ± 0.07	0.028 ± 0.036
Dice loss	0.90 ± 0.08	0.35 ± 0.23	0.91 ± 0.07	0.047 ± 0.031

This suggested that the weighted cross-entropy could enhance the performance of networks and hence improve the accuracy of the segmentation.

4.3.3. Segmentation ablation study

The presented segmentation method comprised three main components of cell detection, cytoplasm segmentation and post-processing. The improved DeepLab model was performed to separate cytoplasm from the image background in a semantic image segmentation task. A data augmentation method was employed to enlarge the labelled samples when training the model. Two post-processing methods (CRFs and DRLSE) were applied after running DeepLab model to further improve the final segmentation results. In order to evaluate the effectiveness of each component in the cell segmentation, an ablation study was conducted on the three test datasets. Table 5 shows the segmentation performance under different ablation configurations.

In this work, a synthesis algorithm [17] was utilized to generate a set of new clustered cell images for the in-house dataset during DeepLab training. Since both ISBI2014 and ISBI2015 datasets used the synthetic images for training, the data augmentation did not applied to these two datasets. We computed four measures, including DSC, FNRo, TPRp, and FPRp to quantify the segmentation performance on the in-house dataset without and with the usage of data augmentation. The quantitative results are shown in Table 5. It can be seen that the data augmentation could be helpful to improve the cell segmentation. Moreover, various combinations of two key parameters (*NUM* and *OR*) were also considered when synthesizing cell images, the changes of parameters had little effect on the segmentation results.

From Table 5, we noted that the cell detection could enhance the cell segmentation with up to 10% DSC improvement on the in-house dataset. This is because that the existence of touching clustered cells in cytology images makes it quite hard for automated segmentation methods to separate them into individual ones. Our two-stage cell detection method using the TernausNet model and the double-window based cell localization strategy was able to accurately identify and determine the individual cell regions, thus serving as the prior knowledge to leverage the performance of the DeepLab model. Also, it can be seen that adding the fully connected CRFs and DRLSE based boundary refinement could help achieve more accurate segmentation than using only CRFs or DRLSE. This was due to the fact that the variations of slide preparation and image acquisition could cause deformation, artifacts and inconsistency of cell appearance, which might impede the segmentation process as well. Hence, the post-processing methods were able to further booster the cell segmentation by refining the outer contours of touching clustered cells with more precise and distinct boundaries.

Moreover, we replaced our network model with the other two popular segmentation models of the U-Net [25] and the fully convolutional networks (FCNs) [39] to evaluate the effective usage of the modified DeepLab model. The U-Net model is a U-shaped fully CNN and has been reported to work efficiently with less annotated training samples. The FCNs is a trained end-to-end pixel-wise semantic segmentation method, which has been demonstrated good performance in segmenting natural images. The quantitative measurement is shown in Table 5. The improved DeepLab model outperformed the other two networks in terms of all the performance metrics, indicating that the developed segmentation method could

Table 5
Segmentation performance under different ablation configurations on three datasets, and the values are in the format of $\mu \pm \sigma$. The bold text indicates the best performance.

loss function	DSC	FNRo	TPRp	FPRp
ISBI2014				
DeepLab	0.85 ± 0.10	0.35 ± 0.33	0.88 ± 0.07	0.005 ± 0.004
DeepLab+Detection	0.90 ± 0.04	0.21 ± 0.26	0.90 ± 0.05	0.002 ± 0.003
DeepLab[a]+CRFs	0.92 ± 0.05	0.13 ± 0.11	0.92 ± 0.04	0.002 ± 0.002
DeepLab[a]+DRLSE	0.91 ± 0.06	0.19 ± 0.16	0.92 ± 0.05	0.002 ± 0.002
DeepLab[a]+CRFs+DRLSE	**0.93** ± 0.04	**0.11** ± 0.13	**0.93** ± 0.05	0.001 ± 0.002
U-Net[a]+CRFs+DRLSE	0.92 ± 0.05	0.25 ± 0.31	0.91 ± 0.05	0.001 ± 0.003
FCNs[a]+CRFs+DRLSE	0.89 ± 0.06	0.29 ± 0.22	0.90 ± 0.06	0.002 ± 0.003
ISBI2015				
DeepLab	0.88 ± 0.07	0.40 ± 0.33	0.87 ± 0.07	0.010 ± 0.013
DeepLab+Detection	0.90 ± 0.05	0.26 ± 0.17	0.89 ± 0.05	0.004 ± 0.004
DeepLab[a]+CRFs	0.92 ± 0.06	0.18 ± 0.13	0.90 ± 0.06	0.003 ± 0.003
DeepLab[a]+DRLSE	0.91 ± 0.06	0.21 ± 0.24	0.88 ± 0.05	0.003 ± 0.004
DeepLab[a]+CRFs+DRLSE	0.92 ± 0.05	**0.13** ± 0.15	**0.91** ± 0.05	**0.001** ± 0.003
U-Net[a]+CRFs+DRLSE	0.90 ± 0.04	0.19 ± 0.14	0.90 ± 0.06	0.002 ± 0.003
FCNs[a]+CRFs+DRLSE	0.88 ± 0.06	0.29 ± 0.25	0.89 ± 0.08	0.003 ± 0.003
In-house dataset				
DeepLab	0.82 ± 0.11	0.46 ± 0.29	0.83 ± 0.09	0.062 ± 0.033
DeepLab+Augmentation	0.83 ± 0.08	0.34 ± 0.26	0.85 ± 0.12	0.053 ± 0.042
DeepLab+Detection	0.90 ± 0.07	0.39 ± 0.43	0.91 ± 0.05	0.026 ± 0.023
DeepLab[b]+CRFs	0.91 ± 0.05	0.20 ± 0.27	0.92 ± 0.07	0.021 ± 0.017
DeepLab[b]+DRLSE	0.91 ± 0.06	0.26 ± 0.24	0.90 ± 0.08	0.025 ± 0.024
DeepLab[b]+CRFs+DRLSE	**0.92** ± 0.04	**0.15** ± 0.13	**0.93** ± 0.05	**0.013** ± 0.012
U-Net[b]+CRFs+DRLSE	0.85 ± 0.06	0.23 ± 0.17	0.88 ± 0.06	0.037 ± 0.031
FCNs[b]+CRFs+DRLSE	0.82 ± 0.11	0.37 ± 0.37	0.86 ± 0.12	0.049 ± 0.042

[a] using both the segmentation model and cell detection;
[b] using the segmentation model, cell detection, and data augmentation.

provide a promising framework to handle the challenging cases of segmenting overlapped cells in cytological images.

4.4. Computational complexity

The segmentation method was implemented under Matlab R2017a platform and Python 2.7 and executed on a machine with a 3.40 GHz Intel Core i5-7500 CPU and 8 GB RAM. The DCNN models were implemented on the open source deep learning framework Pytorch package using NVIDIA GeForce GTX TITAN X with 12 GB memory and accelerated by CUDA 8.0 and cuDNNv 5.1. Due to the large variations in size and shape of cells, the computational time reported here is the average values for performing the cytoplasm segmentation on ISBI2014 and in-house dataset. These two data cohorts contained the same size of images (512×512 pixels). For a fair comparison, we adopted the original implementations of Nosrati's method [14] and Lu's approach [17] running on the same platform. Our segmentation method required approximately 12.34 s per image on ISBI2014 and 17.67 s per image on in-house dataset. Nosrati's and Lu's methods took 25.71 seconds and 187.33 s on ISBI2014, and 35.69 s and 213.62 s on in-house dataset, respectively. This indicated that our method was able to obtain good cell segmentation meanwhile maintaining higher computational efficiency.

5. Concluding remarks

We presented a DCNN-based method that addressed the challenging problem of segmenting each individual cell from a clump of cervical cells in digitalized cytology images. The devised framework went through three main stages: cell detection, cytoplasm segmentation, and boundary refinement. A combination of the improved DeepLab model and computerized image analysis techniques allowed to provide superior performance in segmenting inhomogeneous cell cytoplasm. The method was tested on three independent datasets in both qualitative and quantitative evaluations. The experimental results showed the best DSC, FNRo, and TPRp of 0.93, 0.11, and 0.93, respectively. Validating our method with the previous state-of-the-art algorithms, showed promising results in addressing this research point. Moreover, the achieved remarkable segmentation results demonstrated that the developed DCNN method had the potential to be used for automatic cervical cancer diagnosis and provides powerful support for routinely precancerous screening. In the future work, we will focus on building a unified framework to simultaneously perform individual nuclei and cytoplasm segmentation to further improve accuracy and efficiency, in which highly overlapped nuclei are considered and detected via a learning based method. In addition, we only assessed the method on normal cervical cytology images by varying cell numbers and their degrees of overlap. Abnormal cervical cells can have large variations in the cell appearance indicative of their disease state. Therefore, further work should also adapt the method to accurately segment both normal and abnormal cells by modifying the DCNN architecture and improving optimization process.

Declaration of competing interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and

direct communications with the office). She is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author and which has been configured to accept email from tao.wan.wan@gmail.com.

References

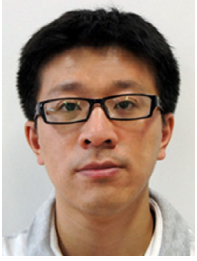
[1] L.A. Torre, F. Bray, R.L. Siegel, J. Ferlay, J. Lortet-Tieulent, A. Jemal, Global cancer statistics, 2012, CA Cancer J. Clin. 65 (2) (2015) 87–108.
[2] W. William, A. Ware, A.H. Basaza-Ejiri, J. Obungoloch, A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images, Comput. Methods Programs Biomed. 164 (2018) 15–22.
[3] Y. Jusman, S.C. Ng, N.A. Osman, Intelligent screening systems for cervical cancer, Sci. World J. 2014 (2014) 1–15.
[4] J. Zhang, Z. Hu, G. Han, X. He, Segmentation of overlapping cells in cervical smears based on spatial relationship and overlapping translucency light transmission model, Pattern Recognit. 60 (2016) 286–295.
[5] N. Ramesh, T. Tasdizen, Cell segmentation using a similarity interface with a multi-task convolutional neural network, IEEE J. Biomed. Health Inform. (2018), (2018). arXiv: http://arxiv.org/abs/10.1109/JBHI.2018.2885544.
[6] J. Song, L. Xiao, Z. Lian, Contour-seed pairs learning-based framework for simultaneously detecting and segmenting various overlapping cells/nuclei in microscopy images, IEEE Trans. Image Process. 27 (12) (2018) 5759–5774.
[7] C. Molnar, I. Jermyn, Z. Kato, V. Rahkama, P. Ostling, P. Mikkonen, V. Pietiainen, P. Horvath, Accurate morphology preserving segmentation of overlapping cells based on active contours, Sci. Rep. 26 (2016) 32412.
[8] M.E. Plissiti, C. Nikou, Overlapping cell nuclei segmentation using a spatially adaptive active physical model, IEEE Trans. Image Process. 21 (11) (2012) 4568–4580.
[9] W. Zhang, H. Li, Automated segmentation of overlapped nuclei using concave point detection and segment grouping, Pattern Recognit. 71 (2017) 349–360.
[10] P. Wang, L. Wang, Y. Li, Q. Song, S. Lv, X. Hu, Automatic cell nuclei segmentation and classification of cervical papsmear images, Biomed. Signal Process. Control 48 (2019) 93–103.
[11] Y. Song, L. Zhang, S. Chen, D. Ni, B. Lei, T. Wang, Graph-based segmentation of abnormal nuclei in cervical cytology, Computer. Med. Imaging Graph. 56 (2017) 38–48.
[12] A. Tareef, Y. Song, H. Huang, D. Feng, M. Chen, Y. Wang, W. Cai, Multi-pass fast watershed for accurate segmentation of overlapping cervical cells, IEEE Trans. Med. Imag. 37 (9) (2018) 2044–2059.
[13] H.A. Phoulady, D. Goldgof, L.O. Hall, P.R. Mouton, A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images, Comput. Med. Imaging Graph. 59 (2017) 38–49.
[14] Z. Lu, G. Carneiro, A.P. Bradley, D. Ushizima, M.S. Nosrati, A.G. Bianchi, C.M. Carneiro, G. Hamarneh, Evaluation of three algorithms for the segmentation of overlapping cervical cells, IEEE J. Biomed. Health Inform. 21 (2) (2017) 441–450.
[15] P. Yan, X. Zhou, M. Shah, S.T.C. Wong, Automatic segmentation of high-throughput RNAi fluorescent cellular images, IEEE Trans. Inf. Technol. Biomed. 12 (1) (2008) 109–117.
[16] A. Tareef, Y. Song, W. Cai, H. Huang, H. Chang, Y. Wang, M. Fulham, D. Feng, M. Chen, Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation, Neurocomputing 221 (2017) 94–107.
[17] Z. Lu, G. Carneiro, A.P. Bradley, An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells, IEEE Trans. Image Process. 24 (4) (2015) 1261–1272.
[18] Y. Song, L. Zhang, S. Chen, D. Ni, B. Lei, T. Wang, Accurate segmentation of cervical cytoplasm and nuclei based on multi-scale convolutional network and graph partitioning, IEEE Trans. Biomed. Eng. 62 (10) (2015) 2421–2433.
[19] L. Zhang, L. Lu, I. Nogues, R.M. Summers, S. Liu, J. Yao, Deeppap: deep convolutional networks for cervical cell classification, IEEE J. Biomed. Health Inform. 21 (6) (2017) 1633–1643.
[20] Y.B. Y. LeCun, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[21] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, F.E. Alsaadi, A survey of deep neural network architectures and their applications, Neurocomputing 234 (19) (2017) 11–26.
[22] Q. Zhu, B. Du, P. Yan, Boundary-weighted domain adaptive neural network for prostate MR image segmentation, CoRR abs/1902.08128 (2019) 1–10,. arXiv: 1902.08128.
[23] Y. Song, E. Tan, X. Jiang, J. Cheng, D. Ni, S. Chen, B. Lei, T. Wang, Accurate cervical cell segmentation from overlapping clumps in pap smear images, IEEE Trans. Med. Imaging 36 (1) (2017) 288–300.
[24] A. Tareef, Y. Song, H. Huang, Y. Wang, D. Feng, M. Chen, W. Cai, Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling, Neurocomputing 248 (2017) 28–40.
[25] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, pp. 234–241.
[26] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2018) 834–848.
[27] G. Thibault, J. Angulo, F. Meyer, Advanced statistical matrics for texture characterization: application to cell classification, IEEE Trans. Biomed. Eng. 61 (3) (2014) 630–637.
[28] T. Wan, J. Cao, J. Chen, Z. Qin, Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features, Neurocomputing 229 (2017) 34–44.
[29] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
[30] C.S. Perone, E. Calabrese, J. Cohen-Adad, Spinal cord gray matter segmentation using deep dilated convolutions, Sci. Rep. 8 (2018) 5966.
[31] Y. Yu, P. Decazes, J. Lapuyade-Lahorgue, I. Gardin, P. Vera, S. Ruan, Semi-automatic lymphoma detection and segmentation using fully conditional random fields, Comput. Med. Imaging Graph. 17 (70) (2018) 1–7.
[32] K. Kamnitsas, C. Ledig, V. Newcombe, J. Simpson, A. Kane, D. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Med. Image Anal. 36 (2017) 61–78.
[33] Z. Zhou, G. Zhao, R. Kijowski, F. Liu, Deep convolutional neural network for segmentation of knee joint anatomy, Magn. Reson. Med. 80 (6) (2018) 2759–2770.
[34] G. Lin, C. Shen, A. van den Hengel, I. Reid, Exploring context with deep structured models for semantic segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 40 (6) (2017) 1352–1366.
[35] A. Arnab, S. Zheng, S. Jayasumana, B. Romera-Paredes, M. Larsson, A. Kirillov, B. Savchynskyy, C. Rother, F. Kahl, P.H. Torr, Conditional random fields meet deep neural networks for semantic segmentation: combining probabilistic graphical models with deep learning for structured prediction, IEEE Signal Process. Mag. 35 (1) (2018) 37–52.
[36] C. Li, C. Xu, C. Gui, M.D. Fox, Distance regularized level set evolution and its application to image segmentation, IEEE Trans. Image Process. 19 (12) (2010) 3243–3254.
[37] M. Rousson, N. Paragios, Shape priors for level set representations, in: Proceedings of the European Conference on Computer Vision (ECCV), 2002, pp. 78–92.
[38] H. Lee, J. Kim, Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and cell-wise contour refinement, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR), 2016, pp. 63–69.
[39] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431–3440.

Tao Wan is an Assistant Professor at Beihang University, Beijing. Dr. Wan was a research associate at the Case Western Reserve University and a postdoctoral associate at the Boston University School of Medicine. She received her Master degree in Global Computing and Multimedia from the Bristol University, UK in 2004 and her Ph.D. in Computer Science from the same university in 2009. She spent one year working as a senior researcher in the Samsung Advanced Institute of Technology (SAIT), China before becoming a visiting scholar in the Visualization and Image Analysis Lab in the Robotics Institute, Carnegie Mellon University. Her research interests are statistical models for image segmentation, fusion, and denoising, machine learning, medical image analysis, computer-aided diagnosis and prognosis system.

Shusong Xu obtained her master degree in biomedical engineering from the Beihang University, Beijing, China. Her research interests mainly focus on image segmentation, deep learning, and computer-aided diagnosis system.

Chen Sang currently is an associate professor in the School of Biological Science and Medical Engineering, Beihang University, Beijing, China. She received her Ph.D from the Nagoya University, Japan, and M.D. from the Bethune Medical University, China. Her research interests are space life science, cellular biology, and molecular biology.

Yulan Jin received her doctor of medicine from the Kochi University, Japan in 2004. She was a visiting scholar in the first department of pathology, Medicine Department of Kochi University before becoming a chief physician in the Beijing Tongren Hospital. She is currently the director of Pathology Department of Beijing Obstetrics and Gynecology Hospital, and an associate professor in the Capital Medical University, Beijing. She has been engaged in pathological diagnosis for 26 years and expertise in diagnostic cytopatholgy.



Zengchang Qin obtained his M.Sc. in Computer Science and Ph.D. in Artificial Intelligence from the University of Bristol, UK, in 2002 and 2005, respectively. He worked as a lecturer in the same university before joining Lotfi Zadeh's BISC group at the EECS Department of UC Berkeley as the BT postdoctoral fellow in 2006. He has been working in Beihang University as an associate professor in the School of Automation Science and Electrical Engineering from 2009. He was also a visiting scholar at Robotics Institute, Carnegie Mellon University, from November 2010 to June 2011. His research interests are uncertainty modeling, machine learning, multimedia retrieval and agent-based modeling.