SEGMENTING OVERLAPPING CERVICAL CELL IN PAP SMEAR IMAGES

Youyi Song, Jie-Zhi Cheng, Dong Ni, Siping Chen, Baiying Lei, Tianfu Wang**

Department of Biomedical Engineering, School of Medicine, Shenzhen University,
National-Regional Key Technology Engineering Laboratory for Medical Ultrasound,
Guangdong Key Laboratory for Biomedical Measurements and Ultra-sound Imaging, Shenzhen, China
{Email: {leiby, tfwang}@szu.edu.cn}

ABSTRACT

Accurate segmentation of cervical cells in Pap smear images is an important task for automatic identification of pre-cancerous changes in the uterine cervix. One of the major segmentation challenges is the overlapping of cytoplasm, which was less addressed by previous studies. In this paper, we propose a learning-based method to tackle the over-lapping issue with robust shape priors by segmenting individual cell in Pap smear images. Specifically, we first define the problem as a discrete labeling task for multiple cells with a suitable cost function. We then use the coarse labeling result to initialize our dynamic multiple-template deformation model for further boundary refinement on each cell. Multiple-scale deep convolutional networks are adopted to learn the diverse cell appearance features. Also, we incorporate high level shape information to guide segmentation where the cells boundary is noisy or lost due to touching and overlapping cells. We evaluate the proposed algorithm on two different datasets, and our comparative experiments demonstrate the promising performance of the proposed method in terms of segmentation accuracy.

Index Terms— Cervical cancer, overlapping cells splitting, multiple-scale deep convolutional networks, dynamic multiple-template deformation model.

1. INTRODUCTION

Cervical cancers developed in the cervical transformation zone are mostly caused by the infection of several types of human papillomavirus. The diseases have a global health impacts second to breast cancer in women. More than 0.28 million women lives were taken by cervical cancer every year, whereas more than 0.53 million women are diagnosed by this type of cancer worldwide [1].

The cytology based screening methods, e.g., the Pap tests, is presently the major approach for detecting cervical cancer. However, the examination of cervical cytological images is manually performed in current clinical practice. As the numbers of cells are quite large, the whole screening process is too time-consuming and extremely labor-intensive. Automated screening methods will be highly desirable to improve the efficiency of screening process and alleviate the manpower shortage problem of medical professionals in the regions with limited medical resources [9,10].

In the literature, most computerized screening methods involved cell segmentation and cell classification. In particular, cell segmentation can help to delimit the cytoplasm and nuclei regions in the stained images. Hence it is a very crucial task for the robust automatic cervical cell analysis.

There were several cervical cell segmentation methods proposed in the literature. In the work [2], cell independence was assumed and promising segmentation accuracy was achieved in local histological image patches. The works [4, 5] further attempted to delineate every cells as many as possible in the full view images. The major problem of theses method lies in the inaccurate segmentation results on the overlapping cells, which can be commonly observed in the stained images, see Fig. 1. The overlapping of cytoplasm is one of the major challenges for the computerized segmentation of cervical cells. The work [3] focused on the splitting of overlapping nuclei for the segmentation of individual cells. In this case, the cytoplasm overlapping issue was indirectly addressed. The approaches [6, 7] employed the shape priors to regularize the segmentation results to demarcate the individual cells in the overlapping cell clumps. These methods [6, 7] impose strong shape prior and regular boundary patterns for the segmentation model, and hence may not be easily generalized to the difficult cases demonstrated in Fig. 1.

By and large, the cell segmentation is a very challenging task due to the complexities of cell structures and highly cell overlapping that lead to poor image contrast (see Fig. 1 a-b). Moreover, other imaging factors like illumination variances, dye concentration caused by the staining procedure, and other variables (i.e., air drying, excessive blood, mucus, bacteria, or inflammation) may also make the automatic visual interpretation more difficult (see Fig. 1 c-e).

To tackle these challenges, we leverage three techniques in this paper to accurately segment individual cervical cell and address the issues of cytoplasm overlapping and ill-defined cell boundaries. Our key contributions are summarized as follows:

- **Cell Component Prediction.** We employ multiple-scale deep convolutional networks (CNN) for feature learning to prediction cell component. Different to traditional CNN based methods, our CNN integrates large contexts into local decisions and allows capturing long-range interactions without training high number of parameters.
- **Overlapping Cells Splitting.** We cast the problem as a discrete labeling task involving multiple cells labels, and

978-1-4799-2349-6/16/$31.00 ©2016 IEEE

(a) (b) (c) (d) (e)

Fig. 1. Illustration of challenges of accurate individual cell segmentation. (a) Complexities of cell structure make accurate segmentation of cytoplasm difficult as shown in violet rectangle. (b-c) Cells highly overlapping and illumination variances lead a poor contrast, dotted lines show related cells' ground truth. (d) Inflammation regions are easily distinguished as cytoplasm. (e) Dye concentration misleads desired segmentation. Besides, as indicated by green arrows, blood, air drying, bacteria and mucus also affect segmentation results.

a suitable cost function is devised and optimized.

- **Cell Boundary Deformation.** The labeling results are used as the dynamic multiple-template deformation model input. The model is guided by the multiple cells' shape. We also explore the cell's structural and contextual information to accurately deform cells' boundaries.
- **Coarse-to-fine Sharp Priors.** We trained shape prior models from coarse to fine levels. The coarse model can serve as good initialization, whereas the fine shape model can assist the refine of segmentation to achieve more accurate segmentation results.

The proposed method is compared with several baseline methods and the experimental results corroborate the efficacy and outperformance of our method.

2. METHODS

A. Feature Learning for Image Representation

We represent images by a probability of background, cytoplasm and nuclei in each pixel. This is important in our method, because we assume each nucleus represents a cell. In addition, the noise of image needs to be transformed into background.

In this paper, we use the multi-scale deep convolutional neural networks to learn image features. The network integrates large contexts into local decisions and allows capturing long-range interactions without introducing high number of extra parameters to train. Specifically, the network contains multiple plane CNN, each plane CNN is applied with different scales what have multiple sizes centered on every pixel in the image. The network can encode more information in a large contextual window and capture appearance, texture, contextual, and shape information when it is properly trained. The structure of our network is illustrated in Fig. 2.

B. Multiple Cells Labeling

The segmentation of overlapping cervical cells in Pap smear images can be defined as a labeling problem. Assuming each image consists of N pixels and L cells (Cells' location and number can be determined by network's output using a simple vote). The goal is to assign a label ℓ from the label set $\mathcal{L} = \{0, 1, \cdots, L\}$ to each pixel i ($y = 0$ represents pixel i belonging to background).

Specifically, let x_i represent the observation of pixel i, the cost of assigning a cell label from $\mathcal{L}$ to every pixel i is

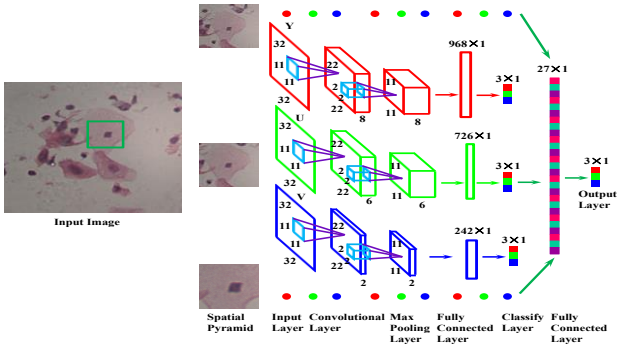

Fig. 2. Structure of the network used for image representation.

defined as:

$$E(\ell; \Theta, \mathrm{w}) = \sum_{i \in \nu} \varphi_i(x_i; \Theta) + \sum_{(i,j) \in \varepsilon} \varphi_{i,j}(x_i, x_j), \quad (1)$$

where ν denotes the set of pixels in image and ε is the neighborhood set of pixel i. w is the parameter vector of the model. The function $\varphi_i(x_i; \Theta)$ is usually called as the unary potential, which denotes the cost of a pixel i taking a label ℓ_i. It is characterized by its shape parameters $\Theta = \{\Theta^1, \Theta^2, \cdots, \Theta^L\}$, where Θ^ℓ denotes the cytoplasm shape for a cell ℓ that is estimated using the Gaussian Kernel (we were fitting the Kernel's parameters that using detected cytoplasm's contour and initializing the Kernel's center with the nuclear centroid). The overlapping-based unary potential is defined as:

$$\varphi_i(x_i = \ell; \Theta) = p_c(i) p^\ell(i) \prod_{\ell, k \in \varepsilon} (1 - p^k(i)), \quad (2)$$

where $p_c(i)$ is the probability of a raw pixel i belonging cytoplasm feeding by the multi-scale deep CNN. $p^\ell(i)$ is the probability of pixel i belonging cell ℓ, which is decided by shape parameter Θ of cell ℓ.

The function $\varphi_{i,j}(x_i, x_j)$ is also known as the pairwise potential, which represents the cost of assigning labels ℓ_i and ℓ_j to two neighboring pixels i and j. This function promotes consistency of labels ℓ_i and ℓ_j ,and is defined as a combination of a smoothness potential $\varphi_s(x_i, x_j)$ and a data dependent edge potential $\varphi_e(x_i, x_j)$:

$$\varphi_{i,j}(x_i, x_j) = \varphi_s(x_i, x_j) + \varphi_e(x_i, x_j). \quad (3)$$

The smoothness potential $\varphi_s(x_i, x_j)$ penalizes dissimilarities between two neighbor pixels and takes the form of an Ising prior:

$$\varphi_s(x_i, x_j) = \begin{cases} 0, & if\ \ell_i = \ell_j \\ \alpha, & otherwise. \end{cases} \quad (4)$$

The edge potential $\varphi_e(x_i, x_j)$ captures the difference of the adjacent pixels' unary potential and is modeled as follows:

$$\varphi_e(x_i, x_j) = 1 - \exp\left(-\frac{\|\varphi_i(x_i) - \varphi_i(x_j)\|}{2\sigma^2}\right), \quad (5)$$

where $\|\cdot\|$ is the Euclidean norm, the parameters α and σ are derived using cross validation on each database.

Our model is defined on a grid structured graph, the graph nodes are corresponding to image pixels. Thus the segmenting problem has been transformed into a labeling problem, and it needs minimizing the energy in Eq. (1) to obtain an optimal labeling $\tilde{\ell} = \arg\min_{\ell} E(\ell; \Theta, \mathrm{w})$. In this paper, we found the most likely labeling is by Maximum a Posterior to estimate $\tilde{\ell}$ upon a set of random variables $\ell \in \mathcal{L}$ by learning the model parameter vector w. The posterior probability w for given N observations X is written as:

$$p(\mathrm{w}|X) = p(\mathrm{w}) \prod_{n=1}^{N} \frac{p(\ell^n|x^n, \mathrm{w})}{p(\ell^n|x^n)}, \quad (6)$$

$$p(\ell|X; \mathrm{w}) = \frac{1}{Z(\mathrm{w})} \exp(-E(\ell; \Theta, \mathrm{w})), \quad (7)$$

where $Z(\mathrm{w}) = \sum_{\ell \in \mathcal{L}} \exp(-E(\ell; \Theta, \mathrm{w}))$ is the partition function. Here we use a zero mean Gaussian prior $p(\mathrm{w})$ distribution over the space of w weight vectors and use Bayes' rule with the independent and identically distributed assumption on observations X. Assuming the likelihood $p(\ell^n|x^n, \mathrm{w})$ is independent of w, the regularized loss function is then defined as:

$$L(\mathrm{w}) = \lambda\|\mathrm{w}\|^2 - \sum_{n=1}^{N} logp(\ell^n|x^n, \mathrm{w}). \quad (8)$$

Finally, we estimate the parameter w by minimizing the loss function using gradient descent techniques. Once the parameter vectors are learned, the inference task is to compute the joint distribution of all the output variables.

C. Dynamic Multiple-Template Deformation

The labeling procedure can accurately label each pixel to corresponding cell. However, pixels in overlapping cells can have multi-labels. A simple example is illustrated in Fig. 3, where the overlapping region is denoted by $S_i \wedge S_j$, and have two labels ℓ_i and ℓ_j. To solve this problem, we follow [6] to perform a joint level set approach. However, we use different energy functions. And instead of using elliptical shapes, we use cases-special templates to guide the evolution process and to determine which reference shape best conform to the evolved boundary.

Specifically, let $\phi: \Omega \to \mathcal{R}$ denotes a level set function (LSF) defined on the image domain, and we have L cells. For split touching clumps, the energy function is minimized as:

$$\mathcal{F}(\{\phi_i\}_{i=1}^{L}) = \sum_{i=1}^{L} \mathbb{E}(\phi_i) + \sum_{i=1}^{L} \sum_{j \in \varepsilon(i)} \psi(\phi_i, \phi_j), \quad (9)$$

where $\varepsilon(i)$ denotes the neighbor cells of cell i. $\mathbb{E}$ is the data energy formimg each cell independently and defined as:

$$\mathbb{E}(\phi_i) = \omega \mathbb{E}_R(\phi_i) + \mathbb{E}_{Ext}(\phi_i), \quad (10)$$

where ω is the weights parameter. The regularization term is defined as:

$$\mathbb{E}_R(\phi_i) \triangleq \int_{\Omega} R(|\nabla \phi_i|) dx. \quad (11)$$

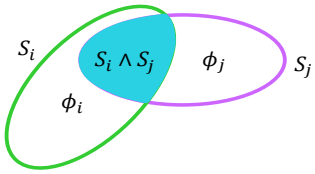

Fig. 3. Illustration of overlapping constraint. Overlapping region belongs to both touching cells, while non-overlapping regions only belong to one cell.

$$R(\phi_i) = \begin{cases} \frac{1}{(2\pi)^2}(1 - \cos(2\pi\phi_i)), & if\ \phi < 1 \\ \frac{1}{2}(\phi_i - 1)^2, & if\ \phi \geq 1. \end{cases} \quad (12)$$

The $\mathbb{E}_{Ext}(\phi_i)$ is an external energy calculated as:

$$\mathbb{E}_{Ext}(\phi_i) = \sum_{l=1}^{2} \lambda_l \int (\int K_\sigma(x - y)|I(y) - \bar{u}_l|^2 M_i(\phi_i(y))\, dy) dx. \quad (13)$$

where $M_1(\phi_i) = H(\phi_i)$ and $M_1(\phi_i) = 1 - H(\phi_i)$, λ_l is a positive constant and H is the Heaviside function. $\bar{u}_1$ and $\bar{u}_2$ are two values that approximate image intensities in ϕ_i with the positive and negative values respectively around x. $I(y)$is the image intensity of pixels located in the region around x,which are acquired by the Gaussian kernel K_σ.

The ψ is a binary functional that defined as:

$$\psi(\phi_i, \phi_j) = \int (H(\bar{s}_i)H(\bar{s}_j) - H(\phi_i)H(\phi_j))\, \varkappa dx, \quad (14)$$

$$\varkappa = \max\left(\bar{s}_i(x), \bar{s}_j(x)\right), \quad (15)$$

where $\bar{s}_i$ is the cases-special template of cell i.

Our cases-special template $\bar{s}_i$ is generated by adaptive hierarchical clustering. Specifically, given two cell's shape s_i and s_j, they are first aligned to the same position by their centroids, then the similarity between s_i and s_j is defined as:

$$\mathbb{s}(s_i, s_j) = 1 - \frac{\mathrm{argmin}_{\theta \in \{1,2,\cdots,\pi\}} \iint_{\Omega} (\phi_{s_i}(\mathcal{S}(x,y)) - \phi_{s_j}(r(x,y)))^2 dxdy}{\iint_{\Omega} (\phi_{s_i}(\mathcal{S}(x,y)))^2 dxdy}. \quad (16)$$

where r is the rotation matrix, $\mathcal{S}$ is the isotropic scaling factor obtained by computing the ratio of mean radius of the shape s_i and s_j. We use an adaptive hierarchical clustering method to cluster over the cells' shape samples. Once the cluster centers are obtained, the cases-special template for each cell S_i is computed by:

$$\bar{s}_i = \mathrm{argmin}_{k \in \{1,2,\cdots,K\}} \mathbb{s}(s_i, s_k). \quad (17)$$

The energy function in Eq. (9) can be optimized by driving the Euler-Lagrange update equation. However, one shortcoming is that the level set approaches are very sensitive to the initialization. Another is that the cervical cells have large natural variability of the shape and the small number of samples available to train, which renders difficult to acquire accurate templates. Therefore, we use labeling results to initialize the level set model and regularly update the templates in the evolution process. As a result, we are able to determine which reference templates best conforms to the evolved boundary best.

3. MATERIALS AND EXPERIMENTS

We tested our method on two different datasets: (1) ISBI 2015 challenge dataset, provided by the ISBI 2015 challenge consisting of 8 cervical cytology images. Each

image has 20~60 Pap stained cervical cells with varying degrees of overlap, and distributed on average in 11 clumps with 3.3 cells per clump. (2) SZU dataset collected from Shenzhen Sixth People's Hospital, Shenzhen, China, consisting of 21 cervical cytology images and prepared using MLBC technique with H&E staining. Each image has 30~80 cervical cells with varying degrees of overlap, and distributed on average in 7 clumps with 6.1 cells per clump.

We evaluate nucleus segmentation using positive predictive value (PPV), negative predictive value (NPV) and F1 measurement (F1). Table 1 compares our method with three state-of-the-art methods [2-4] in nucleus segmentation. It is clear that the performance of our method is better than the selected methods.

To evaluate the cytoplasm segmentation, we use the same measures as ISBI 2015 challenge that includes Dice similarity coefficient (DSC), true positive (TP_P), false positive (FP_P) in the pixel level for well-segmented cells that with DSC>0.7, and false negative (FN_O) in the object level for inaccurate segmentation or missing cells. The evaluation code is obtained from the challenge website. Table 2 summarizes the results of our method and state-of-the-arts including the winners of the ISBI 2015 challenge [8], M. S. Nosrati *et al* [7] and Lu *et al* [6]. From Table 2, we can see that our method outperformed the related approaches in most scenarios. Our method is more significant when the overlapping degrees are higher than the selected methods. Fig. 4 represents some qualitative segmentation results. We can observe that when the number of overlapping cells is around 7, the proposed approach achieves quite good performance even when there are a wide range of cellular size and multiple overlapping types.

Table 1 Quantitative nuclei detection results.

	ISBI 2015			SZU		
	PPV	**NPV**	**F1**	**PPV**	**NPV**	**F1**
[2]	0.64	0.81	0.715	0.52	0.77	0.598
[3]	0.83	0.74	0.782	0.79	0.67	0.710
[4]	0.86	0.89	0.875	0.85	0.90	0.873
Ours	**0.95**	**0.93**	**0.940**	**0.94**	**0.92**	**0.912**

Table 2 Quantitative cytoplasm detection results.

	ISBI 2015				SZU			
	DSC	**TP_P**	**FP_P**	**FN_O**	**DSC**	**TP_P**	**FP_P**	**FN_O**
[8]	0.86	0.90	0.002	0.50	0.78	0.81	0.006	0.47
[7]	0.88	**0.93**	0.005	**0.11**	0.82	0.85	0.006	**0.17**
[6]	0.88	0.92	0.002	0.21	0.81	0.83	0.007	0.27
Ours	**0.89**	0.92	0.002	0.26	**0.84**	**0.88**	**0.004**	0.31

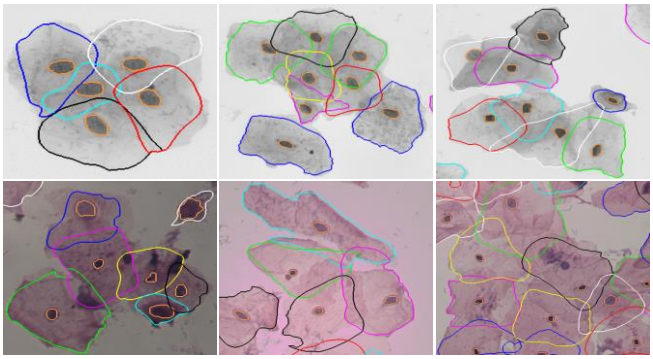

Fig. 4. Qualitative results of overlapping cervical cells segmentation. First row is the segmentation results from ISBI 2015 dataset, while second row represents the segmentation results from SZU dataset.

4. CONCLUSIONS AND FUTURE WORKS

In this paper, we presented a new approach to solve the challenging problem of segmenting individual cell including nucleus and cytoplasm. Our method uses multiple-scale deep convolutional networks to predict pixels' probability, and the overlapping cytoplasm segmenting problems have been transformed into a labeling problem. Also, an overlapping constraint level set is proposed for accurately splitting cell's boundary. We use the cases-special templates to guide the evolution of level set function.

Our future works mainly focus on: (1) Real-time processing of our method using GPU implementation, and (2) Further improvement of the modeling of cells' shape with deep-represented feature, because the appearance cues cannot provide desired results due to the missing or false boundaries.

5. ACKNOWLEDGEMENT

This work was supported partly by National Natural Science Foundation of China (Nos. 61402296, 61571304, 81571758, 61501305 and 61427806), Shenzhen Key Basic Research Project (Nos. JCYJ20150525092940986, JCYJ20150525092940982, JCYJ20130329105033277 and JCYJ20140509172609164), and Shenzhen-HongKong Innovation Circle Funding Program (No. JSE201109150013A).

REFERENCES

[1] WHO/ICO, "Information Centre on HPV and Cervical Cancer (HPV Information Centre), Human Papillomavirus and Related Diseases Report in CHINA," www.who.int/hpvcenter., 2013.

[2] K. Li, Z. Lu, W. Liu, *et al.*, "Cytoplasm and nucleus segmentation in cervical smear images using Radiating GVF Snake," *Patt. Recog.,* vol. 45, no. 4, pp. 1255-1264, 2012.

[3] Y. Al-Kofahi, W. Lassoued, W. Lee, *et al.*, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Trans. Biome. Eng.,* vol. 57, no. 4, pp. 841-852, 2010.

[4] L. Zhang, H. Kong, C. T. Chin, *et al.*, "Automation-Assisted Cervical Cancer Screening in Manual Liquid-Based Cytology With Hematoxylin and Eosin Staining," *Cytom. Part A,* vol. 85, no. 3, pp. 214-230, 2014.

[5] Y. Song, L. Zhang, S. Chen, *et al.*, "Accurate Segmentation of Cervical Cytoplasm and Nuclei Based on Multiscale Convolutional Network and Graph Partitioning," *IEEE Trans. Biomed. Eng.,* vol. 62, no. 10, pp. 2421-2433, 2015.

[6] Z. Lu, G. Carneiro, and A. P. Bradley, "An Improved Joint Optimization of Multiple Level Set Functions for the Segmentation of Overlapping Cervical Cells," *IEEE Trans. Imag. Process,* vol. 24, no. 4, pp. 1261-1272, 2015.

[7] M. S. Nosrati, and G. Hamarneh, "Segmentation of overlapping cervical cells: a variational method with star-shape prior," *in IEEE ISBI*, pp. 186-189, 2015.

[8] G. L. B. Ramalho, D. S. Ferreira, A. G. C. Bianchi, *et al.*, "Cell reconstruction under voronoi and enclosing ellipses from 3D microscopy," *in Overlapping Cervical Cytology Image Segmentation Challenge - IEEE ISBI*, pp. 3-4, 2015.

[9] J.-Z. Cheng, Y.-H. Chou, C.-S. Huang, *et al.*, "Computer-aided US diagnosis of breast lesions by using cell-based contour grouping," *Radiology*, vol. 255, no. 3, pp. 746-754, 2010.

[10] J.-Z. Cheng, Y.-H. Chou, C.-S. Huang, *et al.*, "ACCOMP: augmented cell competition algorithm for breast lesion demarcation in sonography," *Medical physics*, vol. 37, no. 12, pp. 6240-6252, 2010.