Cell Nuclei Segmentation in Fluorescence Microscopy Images Using Inter- and Intra-Region Discriminative Information

Yang Song, *Student Member, IEEE,* Weidong Cai, *Member, IEEE,* David Dagan Feng, *Fellow, IEEE,* and Mei Chen, *Member, IEEE*

***Abstract*— Automated segmentation of cell nuclei in microscopic images is critical to high throughput analysis of the ever increasing amount of data. Although cell nuclei are generally visually distinguishable for human, automated segmentation faces challenges when there is significant intensity inhomogeneity among cell nuclei or in the background. In this paper, we propose an effective method for automated cell nucleus segmentation using a three-step approach. It first obtains an initial segmentation by extracting salient regions in the image, then reduces false positives using inter-region feature discrimination, and finally refines the boundary of the cell nuclei using intra-region contrast information. This method has been evaluated on two publicly available datasets of fluorescence microscopic images with 4009 cells, and has achieved superior performance compared to popular state of the art methods using established metrics.**

I. Introduction

Microscopic image analysis is becoming an enabling technology for modern system-biology research, and cell nucleus segmentation is often the first step in the pipeline. The cell nuclei normally exhibit quite distinctive (higher) intensities from the image background, and they are thus relatively easy for visual identification. However, while many studies have been performed for semi- or fully-automatic nucleus segmentation, the segmentation performance still remains unsatisfactory for more difficult images.

Such difficulties mainly arise from imaging artifacts that: (1) noisy bright regions could appear in the image and could be easily confused as cell nuclei and affect the segmentation of other cell nuclei (e.g. first example in Fig. 1); and (2) the background areas are inhomogeneous and could display higher-than-normal intensities hence becoming harder to differentiate from the adjacent cell nuclei (e.g. second example in Fig. 1). Furthermore, the cell nuclei also appear inhomogeneous that: (1) different cell nuclei could exhibit quite different intensity ranges and hence a global classification criteria might not work well; and (2) the pixels of a single cell nucleus also show varying intensities and some could be very similar to the surrounding background.

Among the numerous nucleus segmentation methods, the morphological methods, such as Otsu thresholding and watershed [1], are widely popular and can be quite effective in cases with good contrast between the cell nuclei and background. More advanced methods, especially those based on level sets [2], [3], are also becoming a major trend for boundary delineation with implicit contour modeling. Another category of methods is focused on pixel-wise labeling accuracy, with various classification algorithms [4], [5], [6] or graphical models [7], [8], [9], [10]. The classification-based methods have the advantage that prior knowledge can be utilized to guide the current segmentation; however, the classification performance is often affected by low separation in the feature space due to image- or nucleus-specific characteristics. The graphical models further incorporate the structural information to enhance the adaptive capability of the algorithm; however, without explicitly addressing the imaging artifacts, such methods might not deliver good performance for more difficult cases.

In this work, we propose an automatic region-based segmentation method for cell nuclei. Considering cell nuclei normally exhibit regular shapes with good intensity contrast from the surrounding background, an initial segmentation is first performed by extract hierarchical salient regions from the image. Then to tackle the imaging artifacts of noisy bright regions and lighter background areas, an inter-region feature discrimination algorithm is designed to reduce the false positives (FP) based on region-level texture and reference-based feature distance. Finally, to better localize the cell nuclei based on the region hierarchy, a graphical model is designed based on intra-region contrast information. Different from classification approaches, the salient region extraction and region-based discrimination are more adaptive to the specific image or nucleus and are thus more effective in handling the inhomogeneity of cell nuclei and background. And with the localized areas generated from the initial segmentation and FP reduction steps, it becomes possible to obtain accurate pixel labeling using a relatively simple graphical model.

II. Methods

Let I represent a microscopic image, which contains N pixels $I = \{p_n : n = 1, ..., N\}$. The objective is to label each pixel p_n as either foreground (i.e. cell nucleus) or background. The labeling of a pixel p_n is denoted as $l_n = \{F, B\}$, and a 3-step approach is designed as following.

A. Salient Region Extraction for Initial Segmentation

In this first step, we extract the salient regions from image I based on the *maximally stable extremal regions* (MSER) [11] algorithm. The MSER method detects regions of regular shapes in hierarchies, with pixels on the region

*This work was supported in part by ARC grants.

Y. Song, W. Cai and D. D. Feng are with the Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, University of Sydney, Australia.

M. Chen is with Intel Science and Technology Center on Embedded Computing, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

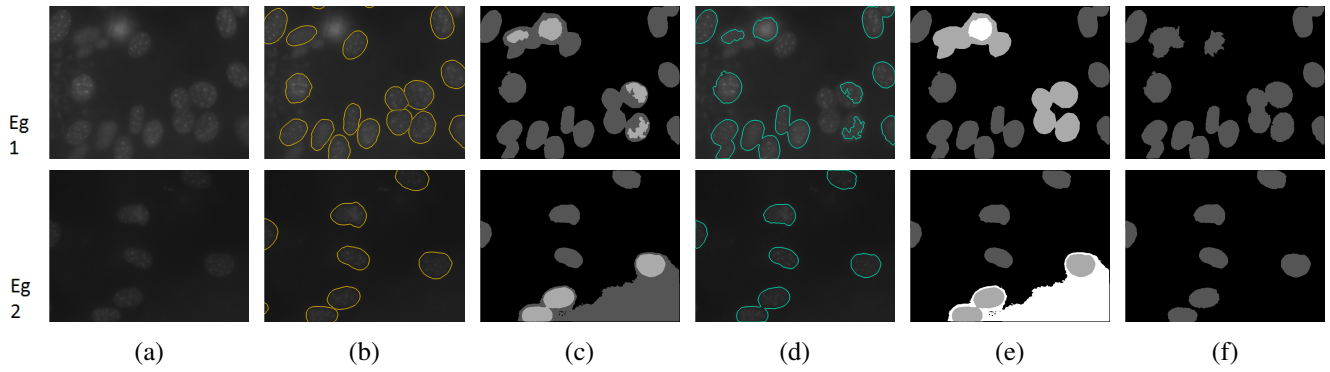


Fig. 1. Illustration of the proposed method flow with two examples. (a) The microscopic image (a quadrant of the original image). (b) Segmentation ground truth indicated as orange contours. (c) Output of initial segmentation, with isolated dark gray areas indicating the single-level regions, and nested light and dark gray areas indicating the top- and bottom-level regions. (d) Regions with green contours represent the reference regions used for FP reduction. (e) Output of FP reduction, with white indicating the FP detected, and light gray indicating the confirmed nucleus regions. (f) Output of final cell localization.

boundary displaying lower intensities than all interior pixels. It is similar to thresholding, but the intensity thresholds are dynamically determined at region-level. The region boundary is also related to level sets, but it does not require any initial contours. Such properties render the MSER method particularly suitable for quick localization of salient objects (i.e. cell nuclei).

To avoid generating many small noisy regions, the image is firstly smoothed using the *anisotropic diffusion filtering* (ADF) [12], which also helps to preserve the edges. The MSER algorithm is then performed [13] to extract a set of salient regions $\{R_x\}$ for an image I.

As shown in Fig. 1c, the extracted regions exhibit two main characteristics. First, for cell nuclei with rather good contrast from the surrounding background, the regions are normally single-level (isolated dark gray regions) and depict the actual nuclei quite closely. Second, for areas clustered with cell nuclei and noisy bright objects (first example), a two-level hierarchy of nested regions is formed, in which the top-level (light gray regions) delineates the noisy bright objects or the high-intensity portions of the cell nuclei, and the bottom-level (dark gray regions surrounding the light gray ones) encloses the cell nuclei and some adjacent background with similar intensities. A two-level hierarchy can be also seen in areas with elevated background intensities (second example); and the top-level actually represents cell nuclei while the bottom-level shows mainly the background.

B. Inter-Region Discrimination for FP Reduction

In the second step, false positive (FP) salient regions are filtered out. As shown in Fig. 1e, the regions at the top- or bottom-level could be FPs, and could appear quite bright or dark. Therefore, filtering based on hierarchy or intensity levels would not be robust. On the other hand, we observe that the real nucleus regions tend to exhibit inhomogeneous textural patterns, while the FP regions are more smooth varying. Furthermore, although cell nuclei in different images might display different features, those in the same image and spatially near actually appear quite similar; while the FP regions are more distinct from the other cell nuclei in the same image. These two observations thus motivate us to discriminate the nucleus and FP regions based on region-level features describing the texture and within-image feature distances.

Specifically, two types of features are computed for each salient region R_x. First, texture features $f_1(R_x)$ based on *local binary patterns* (LBP) [14] are computed as a histogram of the LBP features $lbp(p_i)$ of all pixels $\{p_i\}$ in R_x:

$$f_1(R_x) = \frac{1}{N(R_x)} \sum_{i}^{N(R_x)} lbp(p_i), \forall p_i \in R_x \qquad (1)$$

where $N(R_x)$ is the number of pixels in R_x. The LBP feature $lbp(p_i)$ is computed rotation-invariantly with radius 1 and 8 immediate neighbors. Since rotation invariant LBP feature has 36 possible unique values (i.e. 0 to 35), $lbp(p_i)$ is then represented as a 36-dimensional feature vector with the corresponding element as 1 and the remaining 35 elements as 0. The sum operation in Eq. (1) thus produces a 36-dimensional histogram of LBP features of all pixels in R_x. The feature vector $f_1(R_x)$ is then normalized by the size of R_x to balance between regions. Here LBP is chosen for its good feature descriptiveness and rotation-invariant property; and only 8 immediate neighbors are used for its simplicity and effectiveness based on our experiments.

Second, the feature distance $f_2(R_x)$ between R_x and reference regions $\{R_{x'}\}$ detected in the same image I is computed:

$$f_2(R_x) = \frac{1}{M} \sum_{x'=1}^{M} \|f_1(R_x) - f_1(R_{x'})\|_2 \qquad (2)$$

Here a salient region is considered a reference region $R_{x'}$ if (1) $R_{x'}$ is in the same image quadrant as R_x; and (2) $R_{x'}$ is a single-level region or at the top-level of the hierarchy. And M is the number of reference regions $\{R_{x'}\}$ present for R_x. The example sets of $\{R_{x'}\}$ are shown in Fig. 1d. To better explain the idea, condition (1) ensures the spatial adjacency of $R_{x'}$ and R_x, since it is observed that true cell nuclei would exhibit similar textures if they are spatially near. And we use image quadrant to constrain the spatial distance

for its simplicity. Condition (2) is imposed since bottom-level regions are usually part of the background, while $R_{x'}$ should represent cell nuclei to be used as references for validating R_x. Although some top-level regions could actually be bright background, they are not identified at this stage and hence cannot be reliably excluded. However, including such regions into $\{R_{x'}\}$ would only introduce minor noise with the average operation in Eq. (2), since most reference regions would be true cell nuclei.

Then, based on the feature vector $f_1(R_x)|f_2(R_x)$, R_x is classified as nucleus or FP region using a binary *support vector machine* (SVM), with linear-kernel and other default parameters [15]. Training is conducted on a small subset (about 1/10) of the dataset. The example outputs of the FP reduction can be seen in Fig. 1e. In this step, the single-level regions (dark gray regions in Fig. 1e) are not processed, since they normally represent true cell nuclei and skipping them at this step helps to reduce the computational complexity.

C. Intra-Region Discrimination for Cell Localization

Lastly, the actual cell nuclei are better delineated from the surrounding background. As shown in Fig. 1e, the salient regions detected often contain a mixture of nuclei and background, especially for those formulated in a two-level hierarchy. Therefore, the objective here is to refine the pixel-level labeling for salient and non-FP regions at the top- or bottom-level of the hierarchy. The idea is that, the actual cell nuclei, generally exhibit overall brighter intensities than the surrounding background. To effectively segment foreground from background, a graphical model based on intra-region contrast information is thus designed for the labeling.

Given a salient region R that needs labeling refinement, first, R is expanded using morphological dilation with a fixed width (i.e. 20 pixels) from its contour to include the areas outside of R. Such an expansion is necessary since the region R itself might contain too little background to describe the regional contrast information. For notation simplicity, the expanded region is still denoted as R. A *conditional random field* (CRF) [16] is then constructed for the labeling L_R:

$$E(L_R|R) = \sum_i \phi(l_i) + \alpha \sum_{i,j} \psi(l_i, l_j) \qquad (3)$$

Here $\phi(l_i)$ represents the unary cost of pixel p_i taking the label $l_i \in \{F, B\}$, $\psi(l_i, l_j)$ is the pairwise cost between neighboring pixels p_i and p_j encouraging spatially consistent labeling for region R, and $\alpha = 0.1$ is an empirical weight factor.

To derive the unary cost $\phi(l_i)$, a *k*-means clustering is first used to cluster the pixels in R into two clusters – foreground and background. We observe that due to the rather obvious contrast within a local context, such a simple clustering approach is quite effective in differentiating these two types of pixels. The probability of pixel p_i representing the foreground $Pr(l_i = F)$ is then computed as:

$$Pr(l_i = F) = \min\{\max\{\frac{I_i - I_{R,B}}{I_{R,F} - I_{R,B}}, 0\}, 1\} \qquad (4)$$

where I_i is the pixel intensity of p_i, and $I_{R,B}$ and $I_{R,F}$ are the cluster centroid of background and foreground in R. And with the probability of p_i representing the background as $Pr(l_i = B) = 1 - Pr(l_i = F)$, the unary cost is thus $\phi(l_i) = 1 - Pr(l_i)$. The pairwise cost $\psi(l_i, l_j)$ is defined following the usual Potts model based on the average intensity differences between neighboring pixels.

Using graph cut [17], the energy function $E(L_R|R)$ is then minimized efficiently to obtain the labeling set L_R. Example outputs of the final nuclei localization are shown in Fig. 1f. Note that to reduce the computational complexity, the labeling could be performed at superpixel level (e.g. with mean-shift clustering), and similar results are observed.

III. Results

The proposed method is evaluated on two publicly available sets of 2D fluorescence microscopic images of U2OS and NIH3T3 cells [1]. The two datasets contain 48 and 49 grayscale images, with 1831 and 2178 annotated cell nuclei, respectively. The second dataset (example images shown in Fig. 1 and the second example in Fig. 2) is more challenging than the first dataset (the first example in Fig. 2), with lower contrast and larger intensity inhomogeneity.

To demonstrate the effectiveness of our method, we have compared with four other methods: (1) Otsu thresholding; (2) watershed; (3) level set with handling of intensity inhomogeneity [18] with watershed outputs as the initial contours; and (4) the state-of-the-art [2] reported for the same datasets. For the convenience of comparison, the same performance metrics used in [2] are used here, including the Dice coefficient for pixel-level labeling, the normalized sum of distances (NSD) and Hausdorff distance (HD) for contour delineation, and the number of false positives (FP) and false negatives (FN) for object-level detection. Except Dice, the other four measures are the lower the better.

As shown in Table I and II, the proposed method exhibits promising performance improvements over the compared methods. Since dataset 1 is relatively easy with good contrast between foreground and background in most cases, all methods achieve similar segmentation outputs. Dataset 2 is much more challenging with many bright background regions and dark cell nuclei, the performance improvement with our method is thus much more prominent. The main contributor comes from the object-level accuracy – while the compared methods tend to produce much higher FNs, the proposed region-based approach manages to discover more true cell nuclei with both low FPs and low FNs. In particular, the salient region extraction is especially effective in images with varying contrast. The inter-region feature-based discrimination helps to reduce FPs caused by bright background. The intra-region contrast-based localization further refines the nucleus contour resulting in better the object-level accuracy.

As shown in Fig. 2, the proposed method achieves better segmentation in some typical cases. In the first example, there are large intensity inhomogeneities within cell nuclei; and the level-set method results in some under-segmentation.

TABLE I

THE SEGMENTATION PERFORMANCE ON DATASET 1.

	Dice	NSD	HD	FP	FN
Otsu	0.89	0.06	17.5	1.9	1.7
Watershed	0.94	0.03	11.2	1.4	6.1
[18]	0.93	0.02	9.6	2.5	0.7
[2]	0.94	0.06	13.3	0.5	3.9
Proposed	0.94	0.04	10.8	1.3	1.1

TABLE II

THE SEGMENTATION PERFORMANCE ON DATASET 2.

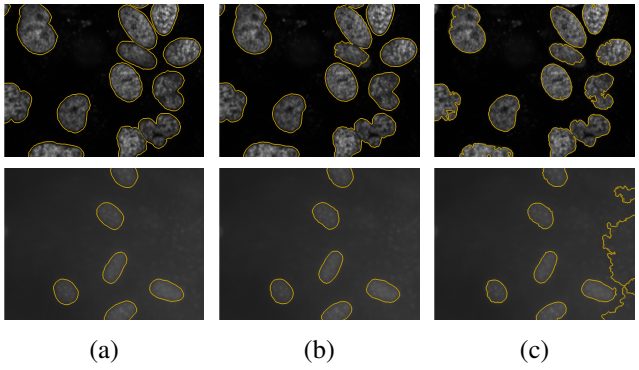
	Dice	NSD	HD	FP	FN
Otsu	0.60	0.42	36.7	1.7	15.3
Watershed	0.78	0.34	47.7	1.8	10.3
[18]	0.80	0.21	25.0	2.2	4.2
[2]	0.83	0.14	16.5	1.7	11.3
Proposed	0.87	0.12	15.8	1.4	2.8

In the second example, due to higher background intensity around the right side of the image, the level-set method could not isolate the actual cell nuclei from the surrounding background. Such results suggest that by modeling localized contrast information, the region-based method can be indeed more effective especially in cases with high inhomogeneities in cell nuclei and background.

It is also worth mentioning that, similar to other pixel labeling-based cell segmentation methods [5], [8], [9], the proposed method does not model cell contours or handle cell splitting. Therefore, cell nuclei that are spatially adjacent or overlapping would merge as one region. However, this study is mainly focused on pixel-wise labeling accuracy, and besides the improved Dice values, the results also show good contour-based measures (NSD and HD) suggesting good approximation of the nucleus boundaries.

IV. CONCLUSIONS

In this work, we proposed a new approach for automatic nucleus segmentation on microscopic images, to effectively handle the intensity inhomogeneity among cell nuclei and background. The proposed method is based on a region-wise formulation, by first creating a hierarchy of salient regions as the initial segmentation. Then based on the set of regions, false positive regions including the noisy bright areas and lighter background areas are filtered out with region-level feature discrimination. The remaining regions in a two-level hierarchy often need further labeling refinement, and a graphical model based on intra-region contrast information is thus designed for final localization. We have evaluated the proposed method on two datasets with altogether 4009 annotated cell nuclei, and shown promising performance improvements over the classic and state-of-the-art approaches.



(a) (b) (c)

Fig. 2. Two example segmentation results. (a) Cell image with segmentation ground truth delineated as orange contours. (b) Segmentation result using the proposed method. (c) Segmentation result using level-set [18].

REFERENCES

[1] L. P. Coelho, A. Shariff, and R. F. Murphy, "Nuclear segmentation in microscope cell images: a hand-segmented dataset and comparison of algorithms," *in Proc. ISBI*, pp. 518–521, 2009.

[2] J. P. Bergeest and K. Rohr, "Fast globally optimal segmentation of cells in fluorescence microscopy images," *in MICCAI LNCS*, pp. 645–652, 2011.

[3] S. Ali and A. Madabhushi, "An integrated region, boundary, shape based active contour for multiple object overlap resolution in histological imagery," *IEEE Trans. Med. Imag.*, pp. 1–14, 2012.

[4] Z. Yin, R. Bise, M. Chen, and T. Kanade, "Cell segmentation in microscopy imagery using a bag of local bayesian classifiers," *in Proc. ISBI*, pp. 125–128, 2010.

[5] L. Cheng, N. Ye, W. Yu, and A. Cheah, "Discriminative segmentation of microscopic cellular images," *in MICCAI LNCS*, pp. 637–644, 2011.

[6] C. Chen, W. Wang, J. A. Ozolek, N. Lages, S. J. Altschuler, L. F. Wu, and G. K. Rohde, "A template matching approach for segmenting microscopy images," *in Proc. ISBI*, pp. 768–771, 2012.

[7] H. Chang, L. A. Loss, P. T. Spellman, A. Borowsky, and B. Parvin, "Batch-invariant nuclear segmentation in whole mount histology sections," *in Proc. ISBI*, pp. 856–859, 2012.

[8] Y. Song, W. Cai, H. Huang, Y. Wang, and D. D. Feng, "Object localization in medical images based on graphical model with contrast and interest-region terms," *in Proc. CVPR Workshop*, pp. 1–7, 2012.

[9] X. Lou, U. Koethe, J. Wittbrodt, and F. A. Hamprecht, "Learning to segment dense cell nuclei with shape prior," *in Proc. CVPR*, pp. 1012–1018, 2012.

[10] Y. Song, W. Cai, and D. D. Feng, "Microscopic image segmentation with two-level enhancement of feature discriminability," *in Proc. DICTA*, pp. 1–6, 2012.

[11] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," *in Proc. BMVC*, pp. 384–393, 2002.

[12] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, 1990.

[13] A. Vedaldi and B. Fulkerson, "Vlfeat: an open and portable library of computer vision algorithms," *in Proc. ACM MM*, pp. 1469–1472, 2010.

[14] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.

[15] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 1–27, 2011.

[16] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: probabilistic models for segmenting and labeling sequence data," *in Proc. ICML*, pp. 282–289, 2001.

[17] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, 2004.

[18] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri," *IEEE Trans. Image Proc.*, vol. 20, no. 7, pp. 2007–2016, 2011.