

# A Multi-Stage Random Forest Classifier for Phase Contrast Cell Segmentation

Ehab Essa<sup>1</sup> and Xianghua Xie<sup>1</sup> and Rachel J Errington<sup>2</sup> and Nick White<sup>3</sup>

**Abstract**— We present a machine learning based approach to automatically detect and segment cells in phase contrast images. The proposed method consists of a multi-stage classification scheme based on random forest (RF) classifier. Both low level and mid level image features are used to determine meaningful cell regions. Pixel-wise RF classification is first carried out to categorize pixels into 4 classes (dark cell, bright cell, halo artifact, and background) and generate a probability map for cell regions. K-means clustering is then applied on the probability map to group similar pixels into candidate cell regions. Finally, cell validation is performed by another RF to verify the candidate cell regions. The proposed method has been tested on U2-OS human osteosarcoma phase contrast images. The experimental results show better performance of the proposed method with precision 92.96% and recall 96.63% compared to a state-of-the-art segmentation technique.

## I. INTRODUCTION

Phase contrast microscopy is an optical based technique that converts difference in phase of object light waves into change in intensity which can be displayed as variations in the image. It produces high contrast images compared to the bright field microscopy of transparent specimens such as living cells directly without need to be killed, fixed, and stained, so that it is used to monitor cell proliferation and examine the drug effect.

Cell segmentation is an important task in order to analyse cell behavior and track its movement across time-lapse images. Manually segmenting cells is a time-consuming, laborious process, that can suffer from high inter- and intra-operator variability, especially in the case of large volume of data captured across time, where each image may contain hundreds of cells. Automatic cell segmentation is still a challenging task, due to low contrast between cell and background, large variations in cell structure, and image artifacts such as halo effect.

Various cell detection and segmentation methods in phase contrast images have been introduced, e.g. thresholding and morphologic operations [1], [2], deformable model based segmentation [3], [4], [5], watershed segmentation [6], [7], [8], graph based methods [9], [10], [11] and machine learning based techniques [6], [12], [13], [2], [14], [15], [8], [16].

<sup>1</sup>E. Essa and X. Xie are with the Department of Computer Science, Swansea University, UK. Correspondence: x.xie@swansea.ac.uk, <http://cvision.swan.ac.uk>. The project is funded by NIS-CHR MIAV-BRU. E. Essa was funded by Swansea University as a PhD student and NISCHR as a postdoc, and he is also affiliated with Mansoura University, Egypt.

<sup>2</sup>R. J. Errington is with the Institute of Cancer & Genetics, School of Medicine, Cardiff University, UK

<sup>3</sup>N. White is with the School of Optometry and Vision Sciences, Cardiff University, UK

Machine learning based methods can be categorized to supervised [6], [12], [13], [2], semi-supervised [14], [8], and unsupervised or clustering based methods [15], [16] depending on the mechanism of the learning. Some of these methods [12], [13], [15] have been used in cell detection based on initial selection of a set of candidate points or small regions corresponding to the cell locations then pruning of less likely candidates using learning-based method. However, these methods are not able to delineate the cells. He *et al.* [6] proposed to use SVM classifier with wavelet features to high-light cell region and seeded watershed method to separate the cell from the background. The seeds are extracted by a separate AdaBoost classifier. In [16], superpixel clustering is used to segment cells based on learning of the cell boundary using a random forest classifier. However, these methods tend to over or under estimate the cell region, particularly when cells form clusters.

Yin *et al.* [2] proposed a phase contrast image restoration method by formulating the problem as a regularized quadratic cost function so that the cell can be segmented by simple thresholding. An SVM classifier is used to identify cells. However, this method is not able to segment bright cells e.g., mitotic cells. Su [14] extends the previous method to segment the bright cells by using different restoration method based on a dictionary representation of diffraction patterns. However user interaction is required to define initial seeds for a semi-supervised method to correctly classify cells.

In this work, we propose a multi-stage RF classification based method to detect and segment cells in microscopy phase contrast images. The first RF classifier is used as a low-level image segmentation to generate a probability map of cell regions. The second RF classifier differentiates the cells from the background noise, and produces delineated cell boundaries.

## II. PROPOSED METHOD

Briefly, the proposed method consists of three steps as shown in Figure 1. First, pixel-wise classification is performed using an RF to generate a probability map of dark and bright cell regions. Next, K-means clustering is used to group pixels into candidate cell regions. Finally, a second RF is introduced to verify the cells and delineate them from the background.

### A. Pixel-wise Classification

Random Forest (RF) [17] is an ensemble classifier from a set of decision trees that can robustly overcome overfitting and outlier problems. RF injects the randomness not only by

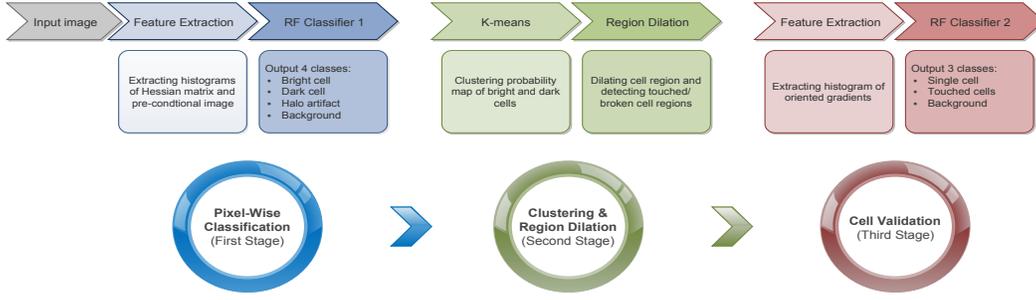


Fig. 1. Proposed system overview.

training each tree on different training sets using a bootstrap sampling but also with a random set of features that is drawn at each node to determine the best tree splitting. In the first stage, we classify image pixels into four categories, i.e. dark cell, bright (mitotic) cell, halo effect, and background.

In RF classification, we train a set of trees  $t \in \{1, \dots, T\}$  independently on subsample of the training data. The testing of unseen data is achieved by taking each testing point  $v$  and pushed it down from the root of all trees until it finds the corresponding leaves. The RF combines the prediction from all trees by simply taking the average prediction of each class  $c$ :

$$P(c|v) = \frac{1}{T} \sum_{t=1}^T p_t(c|v). \quad (1)$$

The setting of RF classifier has two parameters: the number of trees  $T$  and the number of features  $m$  randomly chosen at each split. In this work, we set  $T = 500$  and  $m = 4\sqrt{L}$  where  $L$  is the length of feature vector.

We train RF on two kinds of features, the largest eigenvalue of hessian matrix and the histogram of the pre-conditional features [2], extracted from two sub-windows of size 4 and 8. The output is treated as a probability estimation of the dark and bright cell location.

Hessian matrix has been widely used to enhance vessel and tubular structure [18] and also as ridge detector [19]. Hessian matrix  $H$  is defined as a composition of the second partial derivatives of image  $I$ :

$$H = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix} \quad (2)$$

where  $I_{ij} = \frac{\partial^2 I}{\partial i \partial j}$  and computed by convolving the second derivative of Gaussian filter at scale  $\sigma$  and the image  $I$ . From the obtained Hessian matrix, eigen analysis is performed. Largest eigenvector and its corresponding eigenvalue is a good indicator of the cell location. The scale parameters  $\sigma$  are set as 2 and 4 for each sub-window respectively.

The histogram of the pre-conditioned image is used as an input features for our RF classifier. In this method, a linear model is derived to restore an artifact free image:

$$g = Af \quad (3)$$

where  $g$  is a vectorized observed image,  $A$  is an ill-conditioned matrix determined by the imaging optics, and

$f$  is the restored phase contrast image. Solving the problem by inverting  $A$  is highly noise sensitive. Hence, the problem is formulated as a constrained quadratic function:

$$O(f) = \|Af - g\|_2^2 + \omega_s f^{tr} L f + \omega_r \|\Lambda f\|_1 \quad (4)$$

where  $L$  is a Laplacian matrix working as spatial smoothness regularization,  $\Lambda$  is a positive diagonal matrix in the  $l_1$ -norm sparseness regularization, and  $\omega_s$  and  $\omega_r$  are weighting coefficients. The restored image  $f$  is computed by minimizing the objective function in Eq. (4).

### B. Clustering & Region Dilation

The probability map produced by the first RF classifier based on those low level image features provides an indication of cell location. However, a direct segmentation using the binary output of the classifier is prone to mis-segmentation, particularly when cells form clusters. Instead, we carry out a connected component analysis through spatial clustering and morphological process.

We use the k-means method to find the peak center of the dark and bright cells by clustering the probability values for dark and bright cells. The number of classes is set as 3. We automatically select the output class corresponds to the cells centers by observing the clustering set that maximizes the probability map computed from the first stage. The clustering set is converted to binary label that refer to the potential bright and dark cell regions. Note, the dark and bright cell are treated separately to minimize the overlapping effect.

Cell dilation process is then performed to extend the cell region beyond its center. Let us denote probability maps for dark, bright, and halo artifacts by  $M_d$ ,  $M_b$ , and  $M_h$  respectively. We convert the these probability maps into a binary mask by:

$$B = \begin{cases} 1 & M_d + M_b + M_h > t_b, \\ 0 & otherwise \end{cases} \quad (5)$$

where  $t_b$  is a threshold value and set to 0.55. The result  $B$  represent the likelihood region that can be selected by the dilation process. Each candidate cell is represented by a binary region  $R_i$  in the clustering label (1 for cell, 0 background). The dilated cell region  $C_i$  is computed as:

$$C_i = \begin{cases} 1 & dist(R_i) * B < t_d, \\ 0 & otherwise \end{cases} \quad (6)$$

TABLE I

CELL DETECTION QUANTITATIVE COMPARISON. USING PRECISION, RECALL, F1-SCORE, AND MEAN  $\pm$  STANDARD DEVIATION OF  $D_e$  AND  $D_n$ .

	Precision	Recall	F1-score	$\mu_{D_e} \pm \sigma_{D_e}$	$\mu_{D_n} \pm \sigma_{D_n}$
Proposed Method	92.96	96.63	94.76	$3.16 \pm 2.26$	$10.5 \pm 6.04$
Yin <i>et al.</i> Method [2]	84.08	89.40	86.66	$3.77 \pm 3.25$	$15.6 \pm 10.68$

where  $dist(\cdot)$  is the Euclidean distance transform of the binary region, and  $t_d$  is the distance parameter. Parameter  $t_d$  controls the size of the dilated region, and we set this parameter to 6 for dark cell and 3 for the bright cell. The dilatation process has an advantage that we can determine if the candidate cells centers are touching each other.

### C. Cell Validation

In the final stage, we validate the cell identity by using another RF classifier. We classify the dilated candidate cell regions into three categories, i.e. single cell, touching cells and background. The touching cells detected in the last stage are combined into one region and tested by this classifier to determine if the regions are belonging to the same cell or multiple cells. Histograms of oriented gradients (HoG) [20], and histogram of image intensity are extracted as features from each candidate cell region. HoG is one of the popular image features in object detection. It is based on measuring the distribution of gradient orientations in local regions. The HoG descriptor is implemented by splitting the image into small regions to compute the histogram of edge directions for each region. The histograms of these small regions are concatenated after being normalized within a larger region called block to improve the contrast and illumination invariance. HoG is useful in differentiating single cell from multiple cells connected to each other in that the silhouette of multiple cells are more likely to non-convex and thus their spatial distribution of image gradients differ from that of a single cell. This cell validation stage is essential to remove false positives and minimizing duplicated detections on single cells.

## III. EXPERIMENTAL RESULTS

We test the proposed method on phase contrast images of U2-OS human osteosarcoma cells in control conditions. The time-lapse sequence contains 97 images. Each image in the sequence is  $512 \times 672$  in dimension and cells proliferated from 100+ to 400+ cells. Manual labeling was carried out on every 10th frame (2405 annotated cell) in order to carry out quantitative analysis. The training set includes 10 images (excluded during the testing), where 2 images used to train the first classifier, and 8 images to train the second classifier.

The manual labeling introduced in form of points corresponding to cell centroids. A cell region is considered a true positive (TP) if its centroid is within a small radius distance  $p$  of the ground truth point. If the distance is larger than  $p$ , centroids are considered false positive (FP). The unassigned ground truth points are counted as false negatives (FN). We used 5 evaluation metrics: precision (%), recall (%),  $F_1$ -score (%), Euclidean distance  $D_e$  between

detected cell centroids and the corresponding ground truth points, and finally the absolute difference  $D_n$  between the number of ground truth points and the detected cell centroid numbers. Precision, recall and  $F_1$ -score are defined as:  $Precision = \frac{TP}{TP+FP}$ ,  $Recall = \frac{TP}{TP+FN}$ ,  $F_1 = 2 * Precision * Recall / (Precision + Recall)$ .

Table I shows the quantitative comparison between the proposed method and Yin *et al.* Method [2]. The proposed methods shows higher precision, recall and F1-score and lower distance between cell centroids and ground truth points and also lower absolute difference  $D_n$  in cell numbers compared to Yin *et al.* method. The proposed method can effectively detect both dark and bright cell, in contrast of Yin *et al.* method that can only restore the dark cells.

Figure 2 shows some examples of final segmentation results from the proposed method. The single cell is highlighted by green contour and the touching cell by blue. These touching cells occur when the cells are very close to each other and without clear boundaries in-between. The detection and segmentation of the cells provide the basis for establishing correspondence from one image to another so that the cell growth can be identified and tracked over time.

## IV. CONCLUSION

We presented a machine learning based method to detect and segment the living cells in phase contrast images. A multi-stage RF classifier is proposed to segment and validate the cell regions. K-means clustering and region dilation are used as a link the two classification stages. Cell centers are obtained by K-means clustering from the cell probability map. Region dilation is applied to restore cell region around its center and detect the touching cell. The proposed method showed promising results despite the low contrast and weak edges.

## REFERENCES

- [1] M.A.A. Dewan, M.O. Ahmad, and M.N.S. Swamy, "Tracking biological cells in time-lapse microscopy: An adaptive technique combining motion and topological features," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1637–1647, 2011.
- [2] Zhaozheng Yin, Takeo Kanade, and Mei Chen, "Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation," *Medical Image Analysis*, vol. 16, no. 5, pp. 1047 – 1062, 2012.
- [3] F. Li, X. Zhou, H. Zhao, and S. T. C. Wong, "Cell segmentation using front vector flow guided active contours," in *MICCAI*, 2009, pp. 609–616.
- [4] X. Xie and M. Mirmehdi, "Level-set based geometric colour snake with region support," in *ICIP*, 2003, pp. 153–156.
- [5] I. Seroussi, D. Veikherman, N. Ofer, S. Yehudai-Resheff, and K. Keren, "Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes," *Journal of Microscopy*, vol. 247, no. 2, pp. 137–146, 2012.

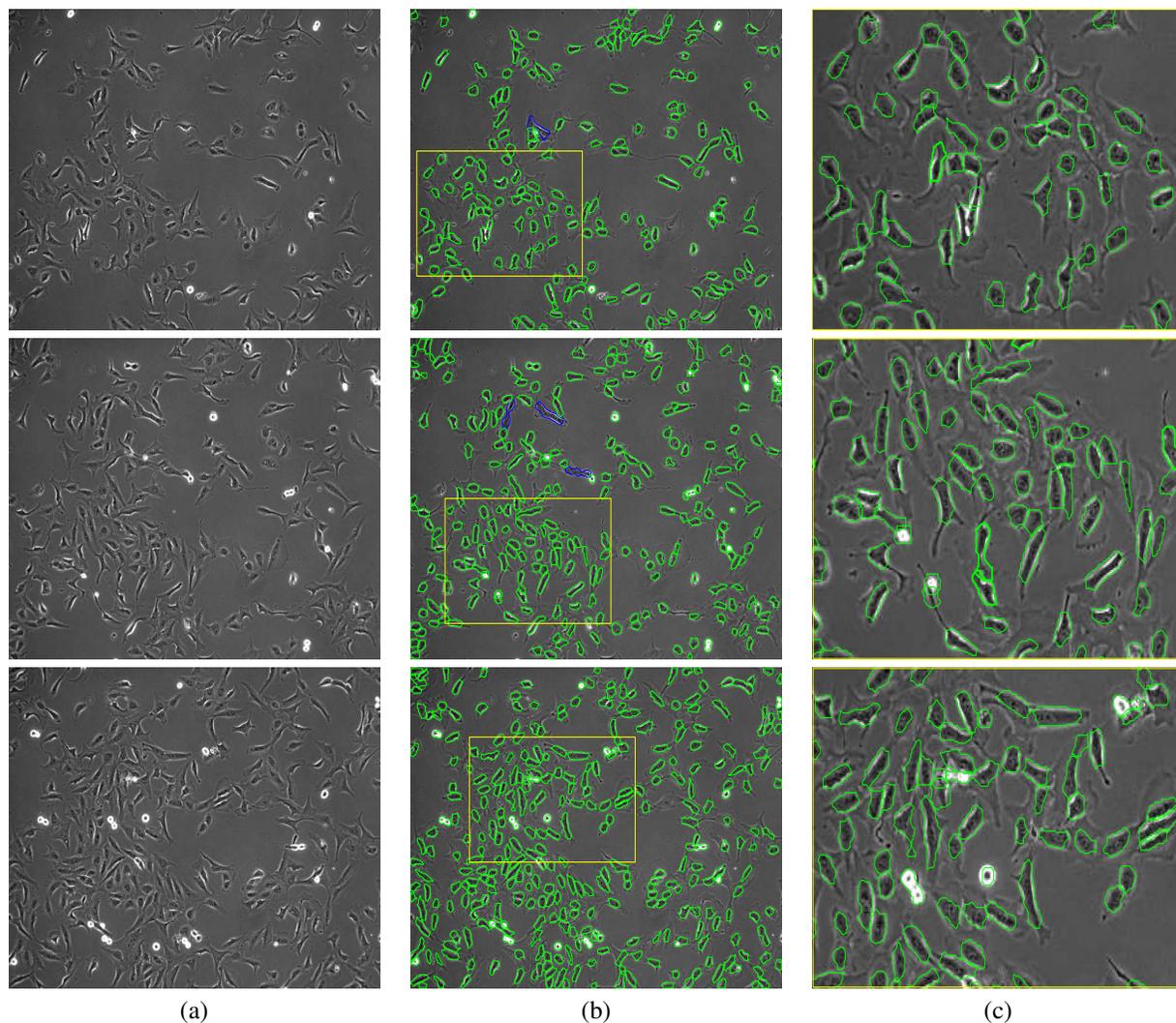


Fig. 2. Proposed cell segmentation results. Cells classified as single (green) or touched (blue): (a) Original image; (b) Result image; (c) Sub-window of the result image.

- [6] Weijun He, Xiaoxu Wang, Dimitris Metaxas, Robin Mathew, and Eileen White, "Cell segmentation for division rate estimation in computerized video time-lapse microscopy," in *SPIE*, 2007, vol. 6431, pp. 1–8.
- [7] O. Debeir, I. Adanja, N. Warzee, P. Van Ham, and C. Decaestecker, "Phase contrast image segmentation by weak watershed transform assembly," in *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2008, pp. 724–727.
- [8] Thomas Nketia, Jens Rittsher, and J. Alison Noble, "Utilizing phase retardation features for segmenting cells in phase contrast microscopy images," in *MIUA*, 2014, pp. 191–196.
- [9] Christopher Russell, D. Metaxas, C. Restif, and P. Torr, "Using the pn pots model with learning methods to segment live cell images," in *ICCV*, Oct 2007, pp. 1–8.
- [10] J. Pan, T. Kanade, and Mei Chen, "Heterogeneous conditional random field: Realizing joint detection and segmentation of cell regions in microscopic images," in *CVPR*, 2010, pp. 2940–2947.
- [11] A. Massoudi, A. Sowmya, K. Mele, and D. Semenovich, "Employing temporal information for cell segmentation using max-flow/min-cut in phase-contrast video microscopy," in *EMBC*, 2011, pp. 5985–5988.
- [12] J. Pan, T. Kanade, and M. Chen, "Learning to detect different types of cells under phase contrast microscopy," in *Microscopic Image Analysis with Applications in Biology Workshop*, 2009.
- [13] Carlos Arteta, Victor Lempitsky, J. Alison Noble, and Andrew Zisserman, "Learning to detect cells using non-overlapping extremal regions," in *MICCAI*, vol. 7510 of *Lecture Notes in Computer Science*, pp. 348–356. 2012.
- [14] Hang Su, Zhaozheng Yin, Seungil Huh, and Takeo Kanade, "Cell segmentation in phase contrast microscopy images via semi-supervised classification over optics-related features," *Medical Image Analysis*, vol. 17, no. 7, pp. 746 – 765, 2013.
- [15] Firas Mualla, Simon Schöll, Björn Sommerfeldt, Andreas Maier, Stefan Steidl, Rainer Buchholz, and Joachim Hornegger, "Unsupervised unstained cell detection by sift keypoint clustering and self-labeling algorithm," in *MICCAI*, vol. 8675 of *Lecture Notes in Computer Science*, pp. 377–384. 2014.
- [16] Chong Zhang, Julian Yarkony, and FredA. Hamprecht, "Cell detection and segmentation using correlation clustering," in *MICCAI*, vol. 8673 of *Lecture Notes in Computer Science*, pp. 9–16. 2014.
- [17] Leo Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [18] AlejandroF. Frangi, WiroJ. Niessen, KoenL. Vincken, and MaxA. Viergever, "Multiscale vessel enhancement filtering," in *MICCAI*, pp. 130–137. 1998.
- [19] Mathews Jacob and Michael Unser, "Design of steerable filters for feature detection using canny-like criteria," *IEEE PAMI*, vol. 26, no. 8, pp. 1007–1019, 2004.
- [20] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *CVPR*, June 2005, vol. 1, pp. 886–893.