DETCIC: Detection of Elongated Touching Cells with Inhomogeneous Illumination Using a Stack of Conditional Random Fields

A. Memariani[1], C. Nikou[1,2], B.T. Endres[3], E. Bassères[3], K.W. Garey[3], and I.A. Kakadiaris[1]
[1]*Computational Biomedicine Lab, Department of Computer Science*
University of Houston, Houston, TX
[2]*Department of Computer Science and Engineering, University of Ioannina, Greece*
[3]*Department of Pharmacy Practice and Translational Research, University of Houston, Houston, TX*
{*amemaria,cnikou,btendres,ebassere,kgarey,ikakadia*}*@central.uh.edu, cnikou@cs.uoi.gr*

Keywords: Cell detection, Conditional random fields, Clostridium difficile infection

Abstract: Automated detection of touching cells in images with inhomogeneous illumination is a challenging problem. A detection framework using a stack of two conditional random fields is proposed to detect touching elongated cells in scanning electron microscopy images with inhomogeneous illumination. The first conditional random field employs shading information to segment the cells where the effect of inhomogeneous illumination is reduced. The second conditional random field estimates the cell walls using their estimated cell wall probability. The method is evaluated using a dataset of *Clostridium difficile* cells. Finally, the method is compared with two region-based cell detection methods, CellDetect and DeTEC, improving the F-score by at least 20%.

1 INTRODUCTION

Developments in scanning electron microscopy (SEM) have facilitated the acquisition of digital images of micron level cells, leading to improvements in cell quantification for pharmaceutical and medical research studies (Endres *et al.*, 2016). However, microscopic images may have inhomogeneous illumination and are often degraded due to noise. Furthermore, the cells have various sizes and are clustered together, making the problem of cell detection challenging.

Recent cell detection methods fall into two categories. The first category assumes that the cells are easily separable from the background. In this family of methods, features are extracted from image patches and are forwarded to a classifier, such as random forests, to identify the cell centroids (Kainz *et al.*, 2015) using several distance metrics for the classification score (Wu and Nevatia, 2009; Wayalun *et al.*, 2012; Saiyod and Wayalun, 2014; Minaee *et al.*, 2014). The second category includes region-based detection methods. At first, cell candidate regions are detected based on shape or statistical texture descriptors. Then, the best candidates are selected based on correlation clustering (Zhang *et al.*, 2014), optimization-based (Arteta *et al.*, 2012; Arteta *et al.*, 2016; Memariani *et al.*, 2016; Browet *et al.*, 2016), or heuristic methods (Keuper *et al.*, 2011; Santamaria-Pang *et al.*, 2015).

DeTEC (Memariani *et al.*, 2016) applied a sequence of two Markov random fields (MRF) to detect touching elongated cells. The first MRF segments the cells from the background using texture features. The second MRF separates the touching cells by estimating the cell walls. However, DeTEC has the following drawbacks: (i) It relies only on texture features and cell wall probabilities to separate cells from their background. Since the algorithm is unsupervised, the features have the same level of importance. However, inhomogeneous illumination may alter the local texture and hence decrease the accuracy of the segmentation. (ii) It applies a number of edge detectors to train a random forest, estimating the cell wall probabilities. However, edge detectors are not robust to noise. In case a cell is eroded due to a laboratory treatment, the random forest detects the erroneous cell walls. (iii) Noisy estimation of cell wall probabilities leads to poor classification of cell walls. (iv) It relies on superpixels; Inhomogeneous illumination hinders the extraction of superpixels whose boundaries match with the cell walls.

Deep neural networks have been applied to microscopy images. DeepCell (Van Valen *et al.*, 2016) applied convolutional neural networks (CNN) to learn the features. However, training a CNN requires a large dataset to tune the parameters and hyperpa-

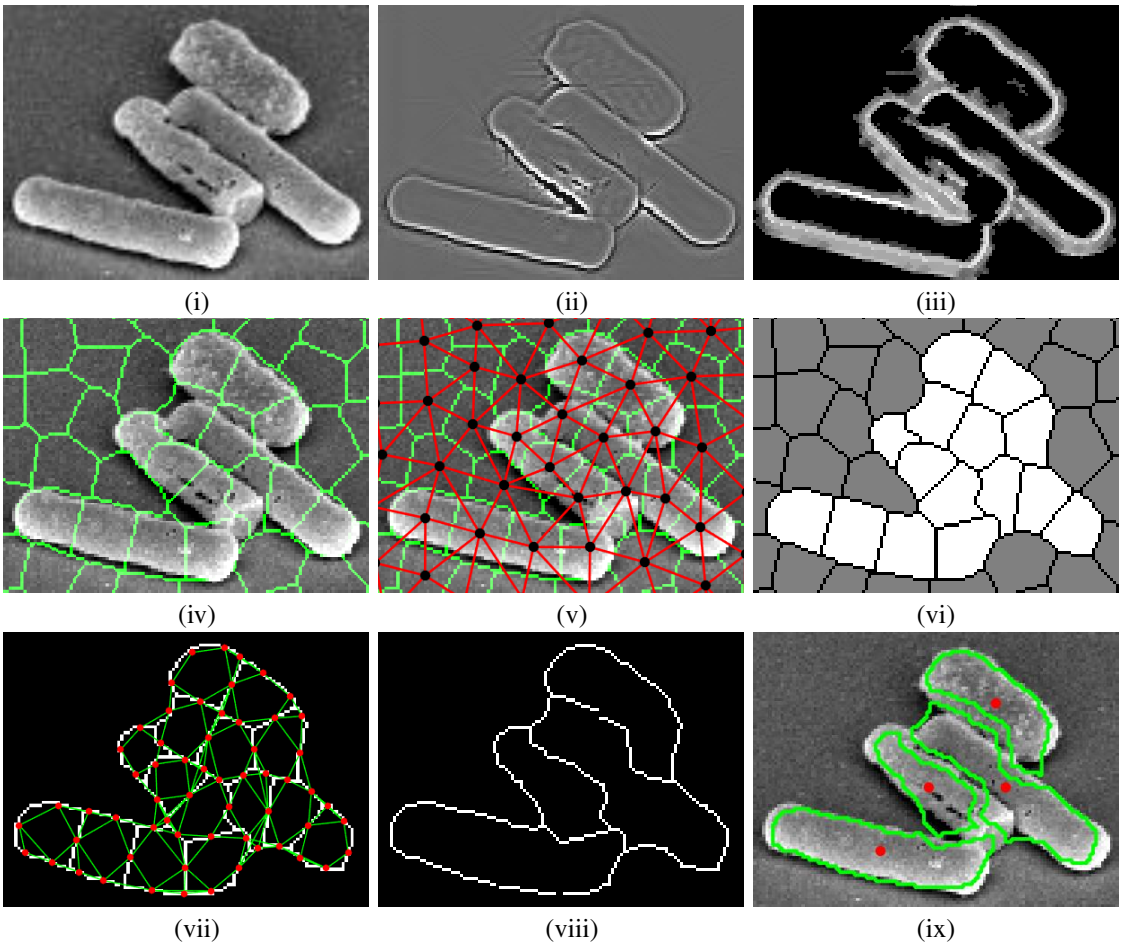


Figure 1: Application of DETCIC to a Clostridium difficile cell image acquired via SEM imaging with 10,000x magnification. (i) Depiction of original image. (ii) Illumination normalization is applied on original image. (iii) A random forest estimates the cell wall probabilities. (iv) Image is divided into superpixels. (v) The first CRF is defined onto superpixels, which segments the cells from their background. (vi) The first CRF segments the cells from their background. (vii) The second CRF is imposed on superpixel boundary components to estimate the cell walls. (viii) The second CRF estimates the cell walls. (ix) Detected cell centroids and their boundaries are shown.

rameters of the networks, which remains a challenge specifically for images obtained by scanning electron microscopy (SEM).

In this paper, DETCIC a detector of elongated cells is proposed which improves the performance of DeTEC with respect to the drawbacks mentioned above. Specifically, (i) DETCIC considers shading along with texture for feature extraction. (ii) it employs a shearlet based edge detector (King *et al.*, 2015) that is robust to noise to enhance the detection of the cell wall pixels. (iii) DETCIC applies a stack of two conditional random fields, which is a supervised method, in contrast to the MRF formulation of DeTEC. (iv) DETCIC applies illumination normalization, reducing the effect of inhomogeneous illumination.

The rest of the paper is organized as follows. Section 2 describes the proposed algorithm. Experimental results are presented in Section 3, comparing the performance of DETCIC with the state-of-the-art cell detection methods DeTEC and CellDetect (Arteta *et al.*, 2012). Finally, conclusions are drawn in Section 4.

2 Methods

DETCIC consists of a stack of two conditional random fields (CRF): the first CRF selects the cell candidates from the background while the second

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

Figure 2: Depiction of edge detector features used for estimation of cell wall probabilities: (a) Original image, (b) Difference of Gaussians, (c) Application of a vessel enhancement filter (Frangi *et al.*, 1998), (d) Roberts edge detector, and (e) A shearlet-based edge detector (King *et al.*, 2015).

CRF separates the touching cells. Estimating the cell walls is an important step for both CRFs. Figure 1 depits the steps of the algorithm. This section describes how the cell walls can be estimated and how the cell wall probabilities can be applied to form the potentials of the two CRFs.

2.1 Estimation of the cell walls

Inhomogeneous illumination hampers the detection of the cell walls. The illumination component is estimated by smoothing the original image in the logarithmic domain using a Gaussian filter. Then, the illumination normalized image is obtained by dividing the image intensities with the estimated illumination in every image $\mathbf{I}$:

$$\mathbf{I}^n = \exp\left(\log\left(\mathbf{I}+1\right) - \log\left(\mathbf{I}+1\right) * \mathbf{G}\right), \qquad (1)$$

where, $\mathbf{G}$ is a Gaussian filter with standard deviation σ_G. The underlying assumption in Eq (1) is the Retinex model (Zosso *et al.*, 2013) of illumination which states that an acquired image $\mathbf{I}$ is a pointwise product of illumination and reflectance. The illumination component is present mainly in coarse scales and it can be estimated by appropriately smoothing the image. The reflectance component captures structures lying, in general, in finer scales.

The illumination normalization highlights the cell walls, reducing the effect of inhomogeneous illumination. A shearlet-based total variation method is applied to obtain the denoised image $\mathbf{D}$, retaining the cell boundaries (Easley *et al.*, 2009).

A random forest estimates the probability of a pixel belonging to a cell wall in $\mathbf{D}$. We compute a matrix of edge detector features $\mathbf{F}^r$, including, difference of Gaussian, a vessel enhancement filter (Frangi *et al.*, 1998), Roberts, and a shearlet-based edge detectors (King *et al.*, 2015). The first two edge detectors are selected because they create a narrow line for cell walls though they may include some noise. On the contrary, the last two features preserve the edges, which have the shape of a curve, but they cover a thicker area around the actual cell walls (Figure 2). The random forest combines all the edge detectors to provide robust boundaries representing the cell walls.

Next, a sequence of two CRFs is describe in which the first CRF finds the cell candidate regions and the second CRF separates cells by estimating their cell walls.

2.2 Cell candidate segmentation

The denoised image $\mathbf{D}$ is divided into superpixels (Mori, 2005). A CRF is applied onto the superpixels with the following objective function:

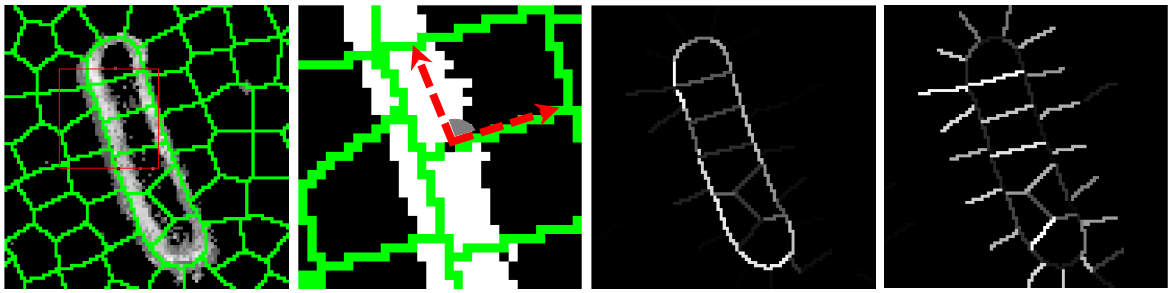
$$E^1 = \sum_t \left(\sum_i u^1_{ti}(\mathbf{f}^1_{ti}, \lambda^1_{ti}; \mathbf{w}^1) + \sum_i \sum_{j \in \mathcal{G}^1_{ti}} v^1_{tij}(\lambda^1_{ti}, \lambda^1_{tj}, \mathbf{P}^1_{tij}; \mathbf{w}^1) \right). \qquad (2)$$

The unary u^1_{ti} and pairwise v^1_{tij} potentials are considered linear in the parameter $\mathbf{w}^1$. The feature vector $\mathbf{f}^1_{ti}$ contains the mean of the shading (Zosso *et al.*, 2013) and intensity values of the i^{th} superpixel.

The pairwise potential v^1_{tij} adds a penalty if the neighboring superpixels have different labels. The pairwise penalty is reduced if the boundary segment between the superpixels i and j has a high probability of belonging to cell wall:

$$\mathbf{P}^1_{tij} = \frac{1}{|\mathcal{N}_{tij}|} \sum_{x \in \mathcal{N}_{tij}} p_x . \cos \alpha_{tij}, \qquad (3)$$

where $\mathcal{N}_{tij}$ is the set of all pixels separating the superpixels i and j in the image t of the training set, and p_x is the probability of a pixel at position x belonging to a cell wall obtained by the trained random



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

Figure 3: (a) Superpixel map (green color) is overlaid onto the cell wall probability map. (b) Zoomed visualization of the area inside the red square in (a). The gray angle is determined by the largest connected component in the probability map (white color) and the superpixel boundary segments (green color). (c) The mean cell wall probabilities of the image depicted in (a). (d) Depiction of the standard deviations of cell wall probabilities.

Algorithm 1: DETCIC training

Input : Training images, cell annotations
Output: Trained random forest, CRF weight parameters

begin
- For every image $\mathbf{I}_t$ $(t = 1,...,n_t)$ in the training set, compute the illumination normalized image $\mathbf{I}_t^n$, shearlet denoised image $\mathbf{D}_t$, superpixel map $\mathbf{S}_t$, and edge detector feature map $\mathbf{F}_t^r$.
- Given the feature map $\mathbf{F}^r$ train a random forest to estimate the cell wall probability $\mathbf{P}^1$
- Given $\mathbf{P}^1$ and $\mathbf{S}_t$ train the first CRF on superpixels, minimizing E^1 to obtain weights $\mathbf{w}^1$
- For every $\mathbf{S}_t$ $(t = 1,...,n_t)$, extract SBCs that belong to a cell wall.
- Train the second CRF on SBCs, minimizing E^2 to learn the weights $\mathbf{w}^2$.

end

forest. The angle α_{tij} is the angle between the superpixel boundary component (SBC) and the corresponding connected component estimated by the random forest when the cell wall probability map is superimposed onto the superpixel map (Figure 3).

The first CRF separates the cell regions from the background by predicting the superpixel labels λ_{ti}^1. However, the cells may be clustered together. Thus, A second CRF is imposed onto the SBCs of the selected superpixels to estimate the cell walls and separate cells.

2.3 Elongated cell separation

The second CRF is defined over the SBCs extracted from the first CRF. The objective function aims to select SBCs that are probable to belong to a cell wall and are elongated:

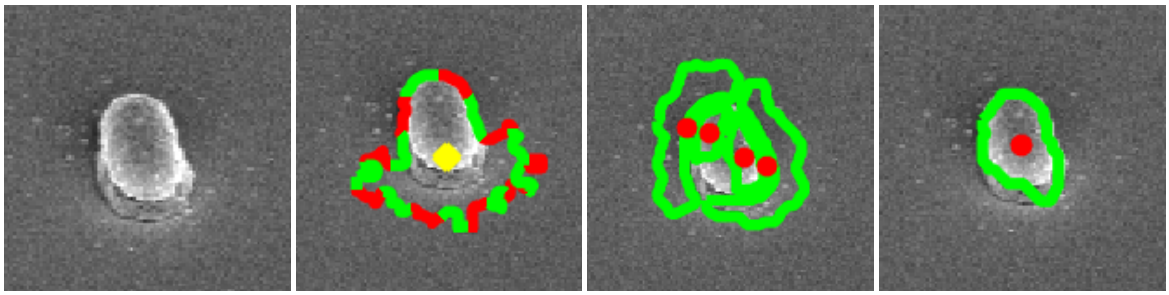
$$E^2 = \sum_t \left(\sum_q u_{tk}^2(\mathbf{f}_{tq}^2, \lambda_q^2, \mathbf{w}^2) + \sum_q \sum_{r \in \mathcal{G}_{tq}^2} v_{tqr}^2(\lambda_q^2, \lambda_r^2, \mathbf{f}_{tq}^2, \mathbf{f}_{tr}^2, \mathbf{B}_{tqr}, \mathbf{w}^2) \right). \quad (4)$$

Similar to the first CRF, the unary and the pairwise terms are linear combinations of features and weight parameters that minimize the energy function E^2. The unary feature vector $\mathbf{f}_{tq}^2$ includes the mean and standard deviation of the cell wall probabilities $\mathbf{P}_{tpq}^2$. The pairwise feature vector includes the difference between the two unary features and the cosine of the angle $\mathbf{B}_{tqr}$ between SBCs q and r. The pairwise potential v_{tqr}^2 penalizes the objective function if the predicted labels λ_q^2 and λ_r^2 are different. However, the penalty is reduced if the two SBCs have different unary features or do not form an elongated structure.

2.4 DETCIC training and inference

The DETCIC training set includes images $\mathbf{I}_t$ $(t = 1,...,n_t)$, which are annotated manually. Cell wall labels to train the random forest are the boundaries of the annotations.

Moreover, the CRF objective function E^1 is trained with the superpixel label set $\mathcal{L}_t^1 = \{l_{ti}^1 \in \{0,1\} | i = 1,...,n_s\}$, where n_s is the number of superpixels in the image. The first CRF



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

Figure 4: Depiction of the effect of inhomogenous illumination: (a) Original image, (b) CellDetect (Arteta *et al.*, 2012), (c) DeTEC (Memariani *et al.*, 2016), and (d) DETCIC.

Algorithm 2: DETCIC inference

Input : A new image $\mathbf{I}_d$, the parameters of the random forest and CRFs
Output: Cell centroids

begin
- For the cell image $\mathbf{I}_d$, compute the illumination normalized image $\mathbf{I}_d^n$, the shearlet denoised image $\mathbf{D}_d$, the superpixel map $\mathbf{S}_d$, and the edge detector features $\mathbf{F}_d^r$
- Input $\mathbf{F}_d^r$ to the trained random forest to compute the cell wall probability map $\mathbf{P}_d$
- Given $\mathbf{P}^1$, $\mathbf{S}_d$, and $\mathbf{w}^1$, apply graph cut to obtain a segmentation on superpixels.
- Extract the SBCs from the selected superpixels.
- Given $\mathbf{P}^2$, and $\mathbf{w}^2$, apply graph cut on SBCs to estimate cell walls.
- Use the estimated cell walls to create morphological connected components.
- Compute the cell centroids.

end

selects superpixels that are likely to belong to a cell. The second CRF is trained with the label set $\mathcal{L}_t^2 = \{l_{tp}^2 \in \{0,1\} | p = 1,...,n_b\}$, where n_b is the number of SBCs extracted from the cell candidate superpixels in the image t in the training set. Label sets $\mathcal{L}_t^1$ and $\mathcal{L}_t^2$ are computed from the manual annotations. Algorithm 1 outlines the training steps for both CRFs. A graph cut provides the labels for each CRF while a gradient-based optimization method selects the best parameter configuration $\mathbf{w}$ that minimizes the objective function E.

Algorithm 1 learns the parameters $(\mathbf{w}^1, \mathbf{w}^2)$. Given a new image I_d, computing the cell wall probabilities $\mathbf{P}_d$ requires computing the illumination normalized image $\mathbf{I}_d^n$ and denoised image $\mathbf{D}_d$ similar to the training images.

Then, DETCIC performs two graph cuts: the first is applied to a rough segmentation of the cells from the background and the second is applied to the SBCs to determine the cell walls (Algorithm 2).

3 Experimental Results

A dataset of *Clostridium difficile* cell images was acquired via SEM imaging with 10,000x magnification and 411×711 pixel resolution. A set of 19 images (211 cells) with similar contrast and cell density were selected for the experiments. The cells are inhomogenously illuminated. Furthermore, cell densities are low in most images but many cells are clustered together, making the detection challenging. In some cases, the cells are partially destroyed due to a laboratory treatment. A GUI is developped for the annotating the cells and the annotations were verified by the expert.

Cell centroids are manually annotated to provide the ground truth. A cell is considered to be detected if the detected centroid lies within a distance d from the ground truth. The distance is set to the length of the smallest cell in the dataset. Precision, recall, and F-score are computed to measure the performance of detection.

Table 1 provides the comparison of the performance of DETCIC with CellDetect and DeTEC. The training was based on a leave-one-out cross validation. CellDetect is a supervised region-based cell detection method which applies extremal regions to detect candidate cell regions (Matas *et al.*, 2004). Then, a statistical model selects the best extremal regions. However, CellDetect fails to detect a fair amount of

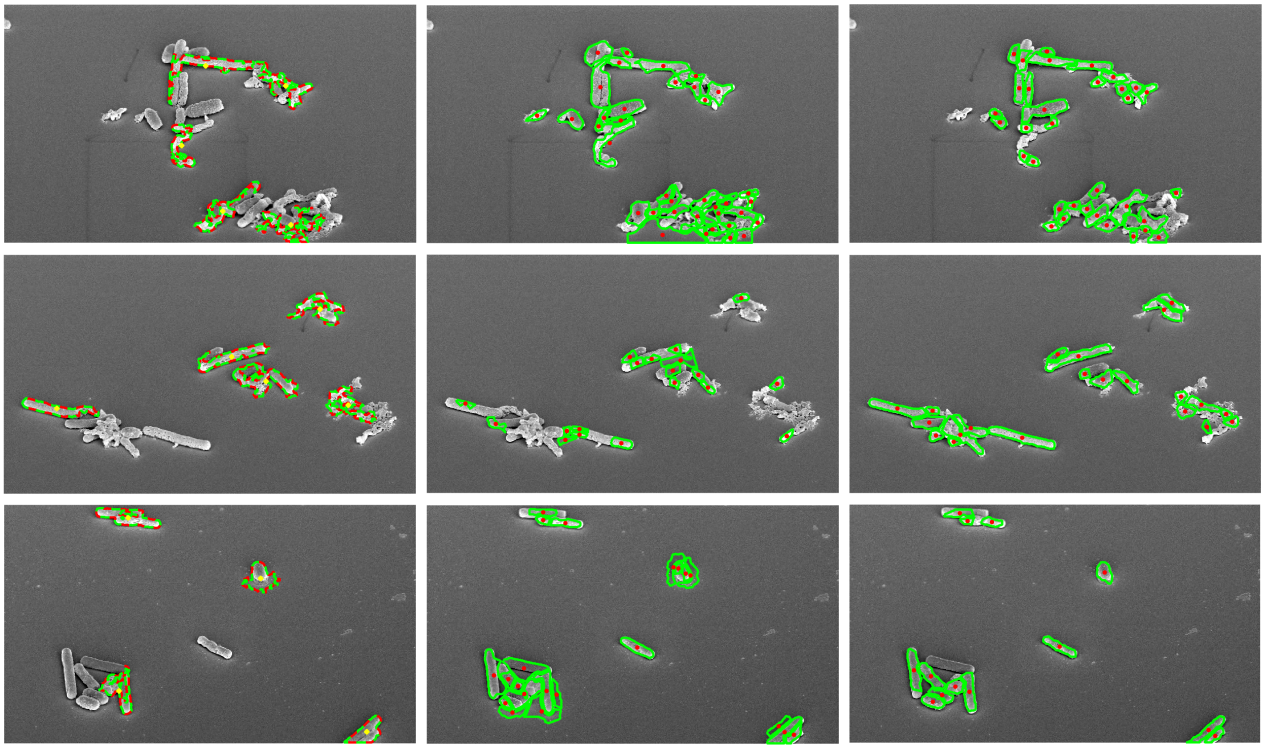


Figure 5: Depiction of the detected cell centroids and their estimated cell walls for CellDetect (Arteta *et al.*, 2012) (Left), DeTEC (Memariani *et al.*, 2016) (Middle), and DETCIC (Right).

Table 1: Comparative results between DETCIC, DeTEC (Memariani *et al.*, 2016), and CellDetect (Arteta *et al.*, 2012), where the acceptable distance of detected centroids from the ground truth is set to the length of the major axis of the smallest cell in the dataset.

Method	Precision	Recall	F-score
CellDetect	**0.80**	0.23	0.36
DeTEC	0.50	**0.88**	0.63
DETCIC	0.68	0.83	**0.75**

cells, assuming there should exist some extremal regions that can represent the cells (Arteta *et al.*, 2016). Therefore, CellDetect achieves a lower recall index compared to the other two methods. DeTEC is an unsupervised region-based method that applies an MRF to segment the cell candidates, and a second MRF to separate the best cell walls to detect the centroids. Although DeTEC detects most cells, the detected cell walls are sensitive to erosion which may be caused by a pharmaceutical treatment. Therefore, some cells are broken into smaller pieces, increasing the number of false positives which leads to low precision. DETCIC significantly improves the cell break downs due to a better estimation of cell wall probabilities which are used to train the second CRF.

Figure 4 depicts an instance where inhomogeneous illumination created shadows on the cell body as well as the area around the cell. CellDetect falsely includes shadows around the cell as part of the cell body. Furthermore, the shadow on the cell body create two bright side on both sides of the cell. DeTEC considers these sides as separate cells and fails to detect the entire cell. However, DETCIC is able to reduce the effect of illumination and detect the cell wall accurately.

Figure 5 depicts examples of detected cells. CellDetect does not detect many cells while failing to separate clusters of touching cells. On the contrary DETCIC is able to detect most cells. However, a few cells are missing due to large shadows which make the cells merge into the background.

DeTEC is able to detect most cells or a portion of them. However, DeTEC fails to estimate the correct boundaries in many cases. Also, DeTEC fails to distinguish between cells and small background regions surrounded by cells due to its unsupervised nature.

Furthermore, DeTEC is more sensitive to inhomogeneous illumination compared to DETCIC. More specifically, DeTEC fails to clearly distinguish be-

tween cells and background in image regions where cell walls are covered by shadows. Figure 4 depicts the detection of a cell effected by inhomogeneous illumination.

4 CONCLUSIONS

A cell detection method (DETCIC) is proposed, that can be used to extract cell meta-data (e.g., number of cells, cell length, cell deformation, etc.). DETCIC is applied on SEM images with inhomogeneous illumination to detect clostridium difficile cells. DETCIC successfully separates touching elongated cells by estimating their cell walls.

ACKNOWLEDEMENTS

This work was supported in part by the National Institutes of Health (NIH/NIAID 1UO1 AI-24290-01) and by the Hugh Roy and Lillie Cranz Cullen Endowment Fund. All statements of facts, opinion or conclusions contained herein are those of the authors and should not be construed as representing official views or policies of the sponsors.

REFERENCES

Arteta, C., Lempitsky, V., Noble, A., and Zisserman, A. (2016). Detecting overlapping instances in microscopy images using extremal region trees. *Medical Image Analysis*, 27:3–16.

Arteta, C., Lempitsky, V., Noble, J. A., and Zisserman, A. (2012). Learning to detect cells using non-overlapping extremal regions. In *Proc. Medical Image Computing and Computer Assisted Intervention*, pages 348–356, Berlin, Heidelberg.

Browet, A., Vleeschouwer, C. D., Jacques, L., Mathiah, N., Saykali, B., and Migeotte, I. (2016). Cell segmentation with random ferns and graph-cuts. In *Proc. IEEE International Conference on Image Processing*, pages 4145–4149, Phonix, AZ.

Easley, G. R., Labate, D., and Colonna, F. (2009). Shearlet-based total variation diffusion for denoising. *IEEE Transactions on Image Processing*, 18(2):260–268.

Endres, B. T., Bassres, E., Memariani, A., Chang, L., Alam, M. J., Vickers, R. J., Kakadiaris, I. A., and Garey, K. W. (2016). A novel method for imaging the pharmacological effects of antibiotic treatment on clostridium difficile. *Anaerobe*, 40:10–14.

Frangi, A., Niessen, W., Vincken, K., and Viergever, M. (1998). Multiscale vessel enhancement filtering. In *Proc. Medical Image Computing and Computer Assisted Intervention*, volume 1496, pages 130–137, Cambridge, MA.

Kainz, P., Urschler, M., Schulter, S., Wohlhart, P., and Lepetit, V. (2015). You should use regression to detect cells. In *Proc. Medical Image Computing and Computer-Assisted Intervention*, pages 276–283, Munich, Germany. Springer.

Keuper, M., Schmidt, T., Rodriguez-Franco, M., Schamel, W., Brox, T., Burkhardt, H., and Ronneberger, O. (2011). Hierarchical markov random fields for mast cell segmentation in electron microscopic recordings. In *Proc. IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 973–978, Chicago, IL.

King, E., Reisenhofer, R., Kiefer, J., Lim, W., Li, Z., and Heygster, G. (2015). Shearlet-based edge detection: flame fronts and tidal flats. In *Proc. SPIE Applications of Digital Image Processing*, volume 9599, pages 1–11, San Diego, CA.

Matas, J., Chum, O., Urban, M., and Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10):761–767.

Memariani, A., Nikou, C., Endres, B. T., Bassères, E., Garey, K. W., and Kakadiaris, I. A. (2016). DeTEC: detection of touching elongated cells in sem images. In *Proc. International Symposium on Visual Computing*, pages 288–297, Las Vegas, NV.

Minaee, S., Fotouhi, M., and Khalaj, B. (2014). A geometric approach to fully automatic chromosome segmentation. In *Proc. IEEE Signal Processing in Medicine and Biology Symposium*, pages 1–6, Philadelphia, PA.

Mori, G. (2005). Guiding model search using segmentation. In *IEEE International Conference on Computer Vision*, pages 1417–1423.

Saiyod, S. and Wayalun, P. (2014). A hybrid technique for overlapped chromosome segmentation of g-band matasprea images automatic. In *Proc. International Conference on Digital Information and Communication Technology and its Applications*, pages 400–404, Bangkok, Thailand.

Santamaria-Pang, A., Rittscher, J., Gerdes, M., and Padfield, D. (2015). Cell segmentation and classification by hierarchical supervised shape ranking. In *Proc. IEEE International Symposium on Biomedical Imaging*, pages 1296–1299, Brooklyn, NY.

Van Valen, D. A., Kudo, T., Lane, K. M., Macklin, D. N., Quach, N. T., DeFelice, M. M., Maayan, I., Tanouchi, Y., Ashley, E. A., and Covert, M. W. (2016). Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLoS Computational Biology*, 12(11):e1005177.

Wayalun, P., Chomphuwiset, P., Laopracha, N., and Wanchanthuek, P. (2012). Images enhancement of g-band chromosome using histogram equalization, OTSU thresholding, morphological dilation and flood fill techniques. In *Proc. 8th International Conference on Computing and Networking Technology*, pages 163–168, Gueongju, China.

Wu, B. and Nevatia, R. (2009). Detection and segmentation of multiple, partially occluded objects by grouping, merging, assigning part detection responses. *International Journal of Computer Vision*, 82:185–204.

Zhang, C., Yarkony, J., and Hamprecht, F. A. (2014). Cell detection and segmentation using correlation clustering. In *Proc. Medical Image Computing and Computer-Assisted Intervention*, pages 9–16, Boston. Springer.

Zosso, D., Tran, G., and Osher, S. (2013). A unifying retinex model based on non-local differential operators. In *Proc. SPIE Computational Imaging*, volume 8657, pages 865702–865702–16, Burlingame, CA.