CELL SEGMENTATION IN MICROSCOPY IMAGERY USING A BAG OF LOCAL BAYESIAN CLASSIFIERS

Zhaozheng Yin[1], Ryoma Bise[1], Mei Chen[2] and Takeo Kanade[1]

[1]Carnegie Mellon University, [2]Intel Labs Pittsburgh

ABSTRACT

Cell segmentation in microscopy imagery is essential for many bioimage applications such as cell tracking and shape analysis. To segment cells from the background accurately, we present a pixel classification approach that is independent of cell type or imaging modality. We train a set of Bayesian classifiers from clustered local training image patches. Each Bayesian classifier is an expert to make decision in its specific domain. The decision from the mixture of experts determines how likely a new pixel is a cell pixel. We demonstrate the effectiveness of this approach on four cell types with diverse morphologies under different microscopy imaging modalities.

Index Terms— Cell segmentation, microscopy image, Bayesian classifier, mixture of experts

1. INTRODUCTION

Noninvasive microscopy imaging techniques such as phase contrast imagery and differential interference contrast (DIC) imagery are suitable to monitor living biological specimens and understand their behaviors. Automated tracking of cell populations in vitro in microscopy enables applications such as optimizing cell culture conditions in stem cell manufacturing to meet research and clinical demands [4]. When developing a computer vision-based tracking system capable of tracking cells in a large population, cell segmentation plays an important role for shape analysis, cell detection and cell association in spatiotemporal context [3, 4, 11].

A simple cell segmentation approach may consist of several steps: compute an intensity histogram of all image pixels, select an intensity threshold based on the statistical histogram, and classify image pixels into two classes based on their intensities and the threshold (pixel class $c = \{C, B\}$, where C and B represent cell and background classes respectively). Because microscopy image histograms may be unimodal, a single thresholding method such as Otsu threshold [6] generates poor results as shown in Fig.1(c). An exceptional work is done by Li and Kanade in [5], where the object image is reconstructed based on a DIC microscopy imaging model and impressive segmentation results are achieved by applying a single global threshold onto the "preconditioned" image.
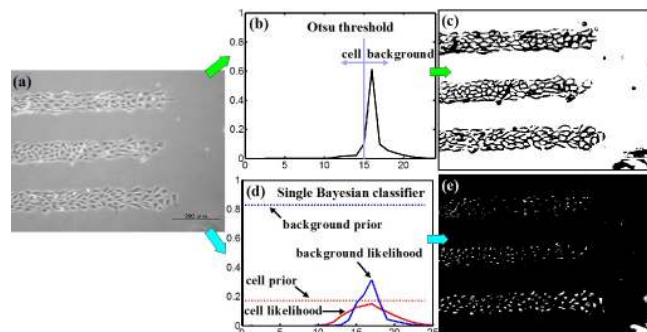

Fig. 1. Cell segmentation is challenging. (a-b-c): segmenting a microscopy image by Otsu threshold [6] yields poor results where many background pixels are misclassified as cell pixels; (a-d-e): segmenting the image by a single global Bayesian classifier also generates bad results. White in (c) and (e) represents cell pixels.

However, this image preconditioning method is computing-intense and it is dependent on the microscopy imaging model.

Some image processing techniques are also widely used in cell segmentation. For example, morphological processing after thresholding an image is used to obtain connected components as cell objects [3], and Laplacian-of-Gaussian (LoG) filters are applied to detect cell blobs [8, 11]. On the blob or object level, efficient features and machine learning techniques are explored to detect cells in microscopy images [7, 12]. In order to segment cells accurately after blob/cell detection, post-processing steps such as level set approaches are usually performed to localize the object boundaries [4, 8].

To achieve accurate cell segmentation with precise boundaries, we consider pixel-level cell classification and formalize it in the basic Bayesian framework. Given features $f(x, \mathbf{I})$ (e.g. intensity, gradient, etc) at pixel x of image $\mathbf{I}$, we classify the pixel as a cell pixel if $P(c = C|f(x, \mathbf{I})) > P(c = B|f(x, \mathbf{I}))$ (maximum a posterior, MAP), where $P(c|f(\cdot))$ is the posterior probability of cell class based on the pixel features. Using Bayes' rule, we have $P(c|f(x, \mathbf{I})) \propto P(f(x, \mathbf{I})|c)P(c)$ where $P(f(\cdot)|c)$ is the likelihood of seeing the feature given the corresponding class, and $P(c)$ is the prior probability about how likely class c will be observed. If we learn the likelihood and prior probabilities from training cell and background pixels all together and build a global Bayesian classifier using intensity features,
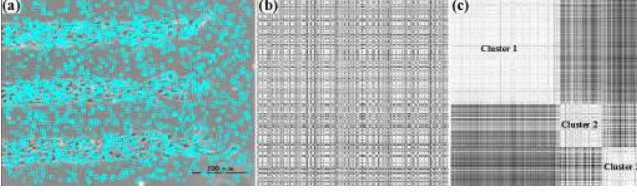
Fig. 2. Cluster local histograms computed around N sample pixels. (a) Around each sample pixel x, we calculate local histograms within its surrounding local windows with width $\mathbf{w}$; (b) we compute a pair-wise similarity matrix among the N local histograms (white represents higher similarity); (c) the N local histograms are clustered into K clusters.

the classification results by MAP are not good (Fig.1(e)). Observing that the likelihood and prior probabilities within a local window vary largely according to different window locations in an image, we propose a cell segmentation approach using a bag of *local* Bayesian classifiers. First, local histograms are computed in training images and clustered into several clusters. From each cluster, we build a Bayesian classifier (expert). The final decision to classify any pixel in a new input image is made by the mixture of experts.

The highlights of this approach include: it is independent of imaging modality and cell type; the segmentation is done by classifying individual pixels with local features, and precise cell boundaries can be achieved in a *soft* way where each pixel is assigned a posterior probability about how likely it is a cell or background pixel. We describe the details of how to build a bag of local Bayesian classifiers in the next section, and demonstrate the effectiveness of this approach in Section 3 with conclusion followed in Section 4.

2. SEGMENTATION USING A BAG OF LOCAL BAYESIAN CLASSIFIERS

A bag of local Bayesian classifiers are combined by a mixture-of-experts model:

$$P(c|f(x,\mathbf{I})) = \sum_{k=1}^{K} \pi_k(x,\mathbf{I},\vec{r}_k) P_k(c|f(x,\mathbf{I})) \qquad (1)$$

where k indexes the K local Bayesian classifiers and $\pi_k(\cdot)$ represents the input-dependent weighting function described in Section 2.2. The intuition behind the model is that different local Bayesian classifiers are responsible to different types of local image patches and they are *experts* to make decision in their expertized domains. The weighting functions determine which expert is dominant based on the input pixel to be classified. In this section, we first describe how to cluster local histograms into K clusters, and then introduce how to build and combine local Bayesian classifiers trained from the clustered image patches.

2.1. Spectral Clustering on Local Histograms

Around any pixel in a training image, we can compute a local histogram within its surrounding local window such as
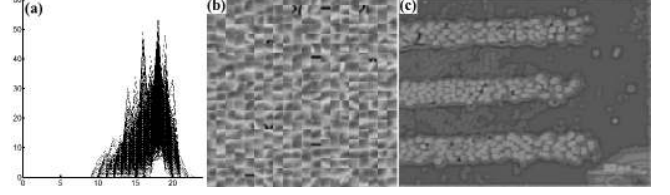

Fig. 3. Learn and combine local Bayesian classifiers. (a) shows a cluster of local histograms; (b) shows some local image patches of the cluster, from which likelihood and prior probabilities are learned; (c) shows the weighting function $\pi_k(x,\mathbf{I},\vec{r}_k)$ of the learned Bayesian classifier on a new image $\mathbf{I}$ to be classified.

(joint) histogram of intensity or (and) gradient magnitude etc. For a microscopy image with 1392×1040 pixels, we can get more than 1.4 millions of local histograms. If learning local Bayesian classifiers from these histograms individually, we would get many similar and redundant classifiers. It is more effective and feasible to cluster local histograms before the learning. During clustering, a pair-wise similarity (or distance) matrix among all histograms is needed. Since a $1.4M \times 1.4M$ similarity matrix can not be stored or operated in a common computer memory, we only compute N local histograms around randomly selected N sample pixels in the training image (e.g. $N = 4000$).

The similarity between any pair of histograms, $S(\vec{h}_i, \vec{h}_j)$, is computed by histogram comparison measures such as Bhattacharya coefficients or Earth Mover's Distance [10]. Fig.2(b) shows an example $N \times N$ similarity matrix among N local histograms. After applying spectral analysis [1] on the similarity matrix, we get K clusters in the permutated similarity matrix as shown in Fig.2(c).

2.2. Learn and Combine Local Bayesian Classifiers

Fig.3(a,b) show a local histogram cluster and its corresponding image patches. We learn local Bayesian classifiers from the clustered image patches. Within each training image patch, we know which pixels belong to cells and which are from the background, then we compute a cell histogram $\vec{h}_k^C$ and a background histogram $\vec{h}_k^B$ over all the pixels of the image patch cluster k. Letting $m_k^C = sum(\vec{h}_k^C)$ denote the number of cell pixels appearing in image patch cluster k, $m_k^B = sum(\vec{h}_k^B)$ be the number of background pixels in the cluster, and $m_k = m_k^C + m_k^B$ be the total number of pixels in the cluster, we compute cluster k's cell prior as $P_k(c = C) = m_k^C/m_k$ and its cell likelihood as $P_k(f(x,\mathbf{I})|c = C) = \vec{h}_k^C/m_k^C$. Using Bayes' rule, we get the cell posterior $p_k(c = C|f(x,\mathbf{I}))$. Similarly, we do the computation for background class. Repeating the process for all clusters, we get a bag of local Bayesian classifiers (cell and background posterior probabilities).

Each local Bayesian classifier is trained from a specific cluster of image patches, thus it is an expert to perform classification on new pixels whose surrounding image patches are
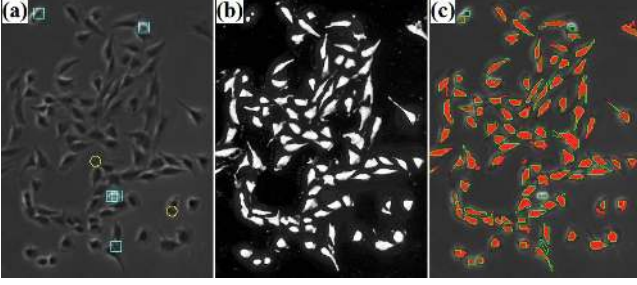
Fig. 4. Experiment evaluation. (a) yellow circle: false alarm, cyan square: miss detection; (b) cell posterior probability $P(c = C|f(x, \mathbf{I}))$, (c) classification results (red mask) by a bag of local Bayesian classifiers match the manual-labeled ground truth (green contours) quite well.

similar to the training image patch cluster. To classify pixel x in a new input image, we need to decide which expert(s) to be applied and what is the confidence level of each classifier's decision, i.e, $\pi_k(\cdot)$ in Eq.1. First, we calculate a local histogram $\vec{h}_x$ around x, and then compute the similarity between $\vec{h}_x$ and every histogram cluster, $S(\vec{h}_x, \vec{r}_k)$, where $\vec{r}_k$ represents the mean histogram of histogram cluster k. The weighting function on classifier k is defined as

$$\pi_k(x, \mathbf{I}, \vec{r}_k) = \frac{S(\vec{h}_x, \vec{r}_k)}{\sum_{i=1}^{K} S(\vec{h}_x, \vec{r}_i)} \qquad (2)$$

which is proportional to the similarity $S(\vec{h}_x, \vec{r}_k)$. For example, the classifier learned from Fig.3(b) is an expert on image patches including many cell pixels, thus in a new input image (Fig.3(c)) the classifier has high confidence around cell pixels and low confidence on other regions. We'd like to point out that the weighting function is *input-dependent* in the mixture model, which is different from the boosting algorithm [2] where the weights are fixed after training and they are proportional to each classifier's accuracy on its training set.

A new input pixel x decides which expert(s) to be applied with what confidence level based on its local histogram. Finally, all experts' decisions are fused into an *ensemble posterior* by Eq.1, and the pixel is classified using its local feature by MAP: $\arg\max_c p(c|f(x, \mathbf{I}))$.

3. RESULTS

The bag of local Bayesian classifier approach is implemented in a scale space. We compute local histograms with several different window sizes (for example, based on the cell sizes in our phase contrast image datasets, we choose $w = \{10, 20, 30\}$). To save computational cost, we use integral histogram technique [9] for histogram computation and Nystrom method for spectral clustering. It costs 50 seconds to classify $1.4M$ pixels on a common desktop workstation, and it can be faster by parallel computing due to the pixel-wise classification.

	Seq1	Seq2	Seq3	Seq4
Precision	0.967	0.898	0.936	0.941
Recall	0.900	0.925	0.854	1.0
F	0.930	0.911	0.893	0.970

Table 1. Quantitative evaluation on four different types of microscopy sequences.

Fig.4(a,b) show a microscopy image to be segmented and its corresponding cell posterior probability after ensembling 9 local Bayesian classifiers (three clusters on each of three window scales). After the pixel-level MAP classification, we group the connected cell pixels into cell candidate blobs and remove small blobs probably due to non-cell pixels. The classification achieves good cell segmentation results compared to manual-labeled ground truth, as shown in Fig.4(c).

We quantitatively evaluate the segmentation on the object level. In the ground truth, a cell object is detected (true positive, TP) if most (e.g. 90%) of its component pixels are correctly classified, otherwise it is missed (false negative, FN). A cell candidate blob by classification is a false positive (FP) if most of its pixels do not match the ground truth. Therefore, we define precision as $P = |TP|/(|TP|+|FP|)$, recall as $R = |TP|/(|TP| + |FN|)$, and F-measure as the Harmonic mean of precision and recall. Our approach is validated on four types of cells of different appearances captured by different imaging modalities and device settings: (1) C2C12 muscle stem cells imaged by Zeiss Axiovert 135TV microscope at 5X magnification; (2) bovine aortic endothelial cells imaged by Olympus IX71 microscope at 10X; (3) bovine vascular endothelial cells imaged at 10X; (4) Central Nervous System (CNS) stem cells imaged by Zeiss Axiovert 135TV microscope at 40X. We use intensity feature on the first three phase contrast microscopy sequences and joint intensity-gradient features on the fourth DIC microscopy sequence. Table 1 shows the complete evaluation results with 92.5% average accuracy (F-measure). Fig.5 shows some samples of the microscopy images and corresponding cell posterior probabilities.

In the end, we perform parameter sensitivity analysis on K over one microscopy sequence. As shown in Fig.6, the segmentation result is poor with only one local Bayesian classifier. After $K \geq 2$, the results are comparable to each other (e.g. we choose $K = 2\sim 5$ in our system).

4. CONCLUSION AND DISCUSSION

We propose a bag of local Bayesian classifier approach for cell segmentation in microscopy imagery. Local Bayesian classifiers (experts) are learned from clustered training image patches. Any new pixel to be classified is assigned a posterior probability about how likely it is a cell or background pixel based on the mixture-of-experts model. The binary segmentation results are obtained by MAP classification. We evaluate our approach quantitatively on four different types of microscopy images with 92.5% average accuracy.
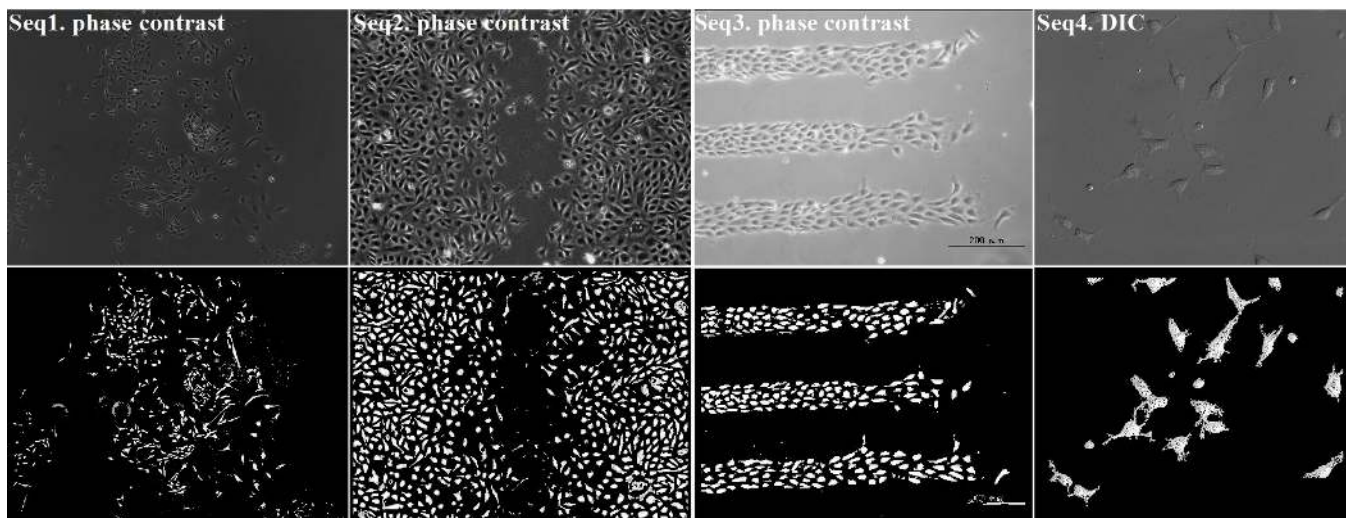
Fig. 5. Sample results on four different microscopy sequences. Top row: input images; bottom row: cell posterior probabilities.
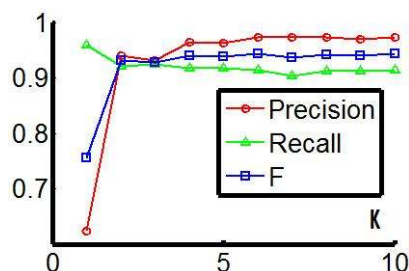

Fig. 6. Quantitative measures on one microscopy sequence (Seq1) with different K's.

Some miss detection happens around the mitosis region as shown in Fig.4(a). The main reason is that we have less training samples around the mitosis region compared to other cell/background regions. In the future, we plan to apply boosting training on these misclassified samples and build a related local Bayesian classifier from them. Another miss detection on the bottom of Fig.4(a) is because we remove the small cell candidate blob, and the two false alarms in Fig.4(a) are due to large cell peripheral parts. The future solution to these problems will explore high-level object classification instead of simply pruning cell candidates by their area sizes.

5. REFERENCES

[1] C. Fowlkes, et al., "Spectral Grouping Using the Nystrom Method," IEEE Transaction on Pattern Analysis and Machine Intelligence, 26(2): p214-225, 2004.

[2] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," International Conference on Machine Learning (ICML), 1996.

[3] D. House et al., "Tracking of Cell Populations to Understand their Spatio-Temporal Behavior in Response to Physical stimuli," IEEE CVPR Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA), 2009.

[4] K. Li, M. Chen and T. Kanade, "Cell Population Tracking and Lineage Construction with Spatiotemporal Context," Intl. Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2007.

[5] K. Li and T. Kanade, "Nonnegative Mixed-Norm Preconditioning for Microscopy Image Segmentation," International Conference on Information Processing in Medical Imaging (IPMI), 2009.

[6] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, 9(1): 62-66, 1979.

[7] J. Pan, T. Kanade and M. Chen, "Learning to Detect Different Types of Cells under Phase Contrast Microscopy," Microscopic Image Analysis with Applications in Biology (MIAAB), 2009.

[8] H. Peng et al., "Integrating Multi-scale Blob/Curvilinear Detection Techniques and Multi-level Sets for Automated Segmentation of Stem Cell Images," IEEE Intl. Symposium on Biomedical Imaging (ISBI), 2009.

[9] F. Porikli, "Integral Histogram: A Fast Way to Extract Histograms in Cartesian Spaces," IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2005.

[10] Y. Rubner, C. Tomasi and L. Guibas, "The Earth Mover's Distance as a Metric for Image Retrieval," Intl. Journal of Computer Vision, 40(2): 99-121, 2000.

[11] K. Smith, A. Carleton and V. Lepetit, "General Constraints for Batch Multiple-Target Tracking Applied to Large-Scale Videomicroscopy," IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2008.

[12] K. Smith, A. Carleton and V. Lepetit, "Fast Ray Features for Learning Irregular Shapes," IEEE International Conference on Computer Vision (ICCV), 2009