

Automated Mitosis Detection of Stem Cell Populations in Phase-Contrast Microscopy Images

Seungil Huh*, Dai Fei Elmer Ker, Ryoma Bise, Mei Chen, and Takeo Kanade

Abstract—Due to the enormous potential and impact that stem cells may have on regenerative medicine, there has been a rapidly growing interest for tools to analyze and characterize the behaviors of these cells *in vitro* in an automated and high throughput fashion. Among these behaviors, mitosis, or cell division, is important since stem cells proliferate and renew themselves through mitosis. However, current automated systems for measuring cell proliferation often require destructive or sacrificial methods of cell manipulation such as cell lysis or *in vitro* staining. In this paper, we propose an effective approach for automated mitosis detection using phase-contrast time-lapse microscopy, which is a nondestructive imaging modality, thereby allowing continuous monitoring of cells in culture. In our approach, we present a probabilistic model for event detection, which can simultaneously 1) identify spatio-temporal patch sequences that contain a mitotic event and 2) localize a birth event, defined as the time and location at which cell division is completed and two daughter cells are born. Our approach significantly outperforms previous approaches in terms of both detection accuracy and computational efficiency, when applied to multipotent C3H10T1/2 mesenchymal and C2C12 myoblastic stem cell populations.

Index Terms—Event detection modeling, mitosis detection, phase-contrast microscopy image analysis, sequential image analysis.

I. INTRODUCTION

STEM cell research has attracted increasing attention due to its enormous potential in regenerative medicine for replacing damaged or diseased tissues or organs. Methods for assessing the proliferative activity of stem cells grown *in vitro* are critical tools for monitoring the health and growth rate of a cell population. Such methods have historically relied on detecting mitosis [1], which is the process whereby the genetic material of a eukaryotic cell is equally distributed between its descendants through nuclear division, resulting in the birth of daughter cells.

Manuscript received September 09, 2010; revised October 13, 2010; accepted October 14, 2010. Date of publication October 21, 2010; date of current version March 02, 2011. This work was supported in part by the National Institutes of Health under Grant RO1EB004343 and Grant RO1EB007369, and in part by the Pennsylvania Infrastructure Technology Alliance (PITA). The work of S. Huh was supported by a Samsung Fellowship. *Asterisk indicates corresponding author.*

*S. Huh is with Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: seungilh@cs.cmu.edu).

R. Bise and T. Kanade are with Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA.

D. F. E. Ker is with the Department of Biological Sciences and Biomedical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA.

M. Chen is with Intel Labs Pittsburgh, Pittsburgh, PA 15213 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2010.2089384

Presently, many cell proliferation assays that are compatible with automated sample handling and high-throughput screening have been developed to measure cell proliferation [2]. However, the majority of these procedures utilize fluorescent, luminescent or colorimetric assays which may require destructive methods of cell manipulation, such as cell lysis and *in vitro* staining, and do not allow for continuous monitoring of cells in culture.

Since phase-contrast microscopy is a form of nondestructive imaging, automated time-lapse systems employing this imaging modality for monitoring and analyzing cell populations *in vitro* have enormous potential for cell biology and stem cell engineering [3]–[6]. Such systems not only enable high-throughput analysis of time-lapse microscopy images, but also facilitate continuous monitoring of live and intact cells for studying biological phenomena and quantifying various cell responses. Moreover, since samples are continuously monitored, sampling at various time points is abrogated, resulting in reduced human labor and eliminating the costs of expensive reagents required for cell lysis or *in vitro* staining. Within the context of such microscopy systems, automated detection of mitosis can provide quantitative information regarding cell proliferation on a continuous basis. This functionality is also expected to improve automated cell tracking systems, a comprehensive tool for the analysis of cell behavior [5], [7], [8].

In order to achieve these benefits, automated mitosis detection systems are required to localize *birth events*; we define a birth event in each mitotic event as the time and location at which the two daughter cells first appear and the boundary between the two is clearly observed. Accurate detection of birth events facilitate the quantification of biological metrics, such as the mitotic index and synchrony, allowing biologists to experimentally assess how altering the conditions under which cells are cultured can impact population growth. In addition, birth event information is helpful in determining when and where a trajectory of a mother cell branches into two trajectories of its daughter cells in cell tracking. Furthermore, precise localization of birth events may aid in the discovery and characterization of novel biological phenomenon that occur at rare frequencies such as a single cell division event giving rise to more than two daughter cells.

Several mitosis detection methods for phase-contrast time-lapse microscopy images have been proposed based on cell tracking. Yang *et al.* [7] obtained blob regions along each cell's trajectory produced by a tracking method. Each blob region is then examined to determine if it contains a mitotic event based on several blob properties, including area, perimeter, circularity, and average intensity. Debeir *et al.* [9] adopted a combination of several mean-shift-processes to track

cells using an ensemble of nested kernels. One of the kernels was designed to model cells in the mitotic state by taking into account their morphological changes. Al-Kofahi *et al.* [10] presented a multiple-object matching method that can handle cell divisions in the typical frame-by-frame segmentation tracking method. Padfield *et al.* [11] investigated cell cycle phases through tracking each nucleus over time. Mitotic events were then identified by linking nonmitotic phases using both the Euclidean distance metric and the fast marching method. These approaches are intuitive but limited in that mitosis detection is dependent on tracking performance, which is generally more challenging to achieve than mitosis detection performance itself. Several mitosis detection algorithms that do not involve cell tracking have recently been proposed. Li *et al.* [5] applied a fast cascade learning framework [13] adopting AdaBoost classifiers [14] to volumetric Haar-like features extracted from spatio-temporal patches covering the whole image region. This approach requires a large number of training samples and searches through the entire space due to the lack of explicit candidate detection. Debeir *et al.* [12] proposed a method to detect mitotic cell regions based on brightness change and link the regions in consecutive frames into a cell division candidate linkage. Each candidate was then validated based on its length (the number of frames). This approach is efficient due to the reduced search space by the candidate detection, but the validation scheme is too simple to effectively distinguish actual mitotic events from the other candidates.

More recently, Liu *et al.* [15] proposed an approach to compensate for the drawbacks of the previous methods. After mitosis candidate patch sequences are constructed through 3-D seeded region growing, Hidden Conditional Random Fields (HCRF) [16] are trained to examine each of the candidates. This approach does not resort to tracking and adopts explicit candidate detection as well as model-based validation, achieving good performance on C3H10T1/2 stem cell populations. However, this work does not detect birth events and is as such limited in that its mitosis detection results may not be sufficient for accurate quantitative analysis of cell proliferation or cell tracking. In fact, the HCRF model is intrinsically not capable of modeling birth events. In addition to inadequate modeling power, this approach is computationally expensive due to the preconditioning step [17], which was originally devised to segment nonmitotic cell regions from background rather than mitotic cell regions.

In this paper, we propose an effective mitosis detection approach that explicitly detects birth events. We developed a probabilistic model that not only determines whether a mitotic event occurs, but also provides the time at which the mitosis is completed and daughter cells are born. By additionally handling the information of cell birth, EDCRF achieves higher discriminating power than HCRF in the identification of mitosis occurrence. In addition, EDCRF is superior to possible alternatives based on HCRF in the temporal localization of birth events as well. We conducted experiments on multipotent C3H10T1/2 mesenchymal and C2C12 myoblastic stem cell populations, and achieved good performance on the population with high confluency. Furthermore, we bypassed the time-consuming preconditioning without sacrificing performance; as a result, computa-

tional efficiency is significantly improved compared to the previous work [15].

The remainder of this paper is organized as follows. We introduce the overall process of our approach and describe the details of the image processing part in Section II. We then formulate a probabilistic model to determine mitosis occurrence and birth event timing in Section III. The experimental setup and results with discussions are presented in Sections IV and V, followed by conclusions in Section VI.

II. APPROACH

Given a sequence of phase-contrast microscopy images, our goal is to detect mitosis in the sequence and localize a birth event during mitosis. To achieve this goal, we adopt the idea of candidate patch sequence construction from the previous work [12] and form a process comprising three steps: candidate patch sequence construction, visual feature extraction, and identification of mitosis occurrence/temporal localization of birth event. We sketch each step to provide an overall view of our approach. The detailed methods will be subsequently described in this and the following sections.

- *Candidate patch sequence construction*: The goal of this step is to detect all spatio-temporal patch sequences that contain mitosis, while detecting as small a number of sequences not containing mitosis as possible. Through this step, mitotic events are spatially localized and the search space is significantly reduced from entire image sequences to candidate patch sequences. As a result, the subsequent steps can be more efficiently conducted, while maintaining mitosis detection accuracy. Fig. 1 shows some examples of candidate patch sequences our method automatically extracted.
- *Visual feature extraction*: Visual features are extracted from each patch of candidate patch sequences based on the characteristics of phase-contrast microscopy images.
- *Identification of mitosis occurrence/Temporal localization of birth event*: For each candidate patch sequence, we perform two decision tasks regarding mitosis occurrence and birth event. The identification of mitosis occurrence determines whether each candidate patch sequence contains a mitosis, specifically, a birth event. For each patch sequence determined to contain a birth event, the temporal localization of the birth event decides which patch contains the birth event in the patch sequence.

A. Candidate Patch Sequence Construction

For preprocessing, we compute the average image of all images in a given sequence; the average image is then subtracted from each image. By this simple procedure, stationary bright artifacts which may be misrecognized as mitosis cell candidates are removed. In addition, intrinsic illumination variation in phase-contrast microscopy images can be corrected.

For each image, candidate patches are first extracted based on brightness (pixel intensity); it is known that the process of mitosis typically exhibits a series of distinctive cell features including increased brightness, increased circularity, and decreased size [5] (see Fig. 2). Each preprocessed image is convolved with a small-sized $d \times d$ rectangular average filter and

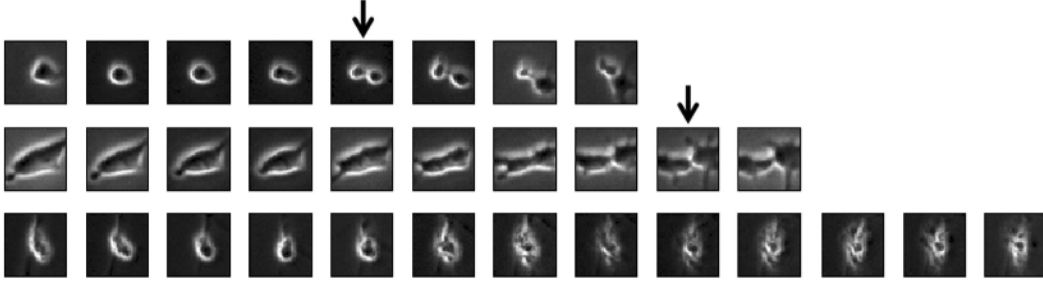


Fig. 1. Positive (top two) and negative (bottom) candidate patch sequences extracted from a multipotent C3H10T1/2 mesenchymal stem cell population. The numbers of patches in candidate patch sequences are varied. The arrows indicate the frames containing manually annotated birth events, the location and time in which daughter cells are born. The end of the mitotic phase involves cytokinesis, which is the pinching of the cell membrane to split a cell into two, typically shown as a figure eight shape as in the first sequence; however, a figure eight shape is often not clearly observed as in the second sequence. The last sequence contains apoptosis, or cell death, which exhibits similar visual characteristics to mitosis in the beginning of the process. Due to the increased brightness, apoptotic events are mostly captured in candidate patch sequences. These sequences are extracted from frames 86 through 93 (top), 350 through 359 (center), and 819 through 831 (bottom) among the entire sequence consisting of 1436 frames.

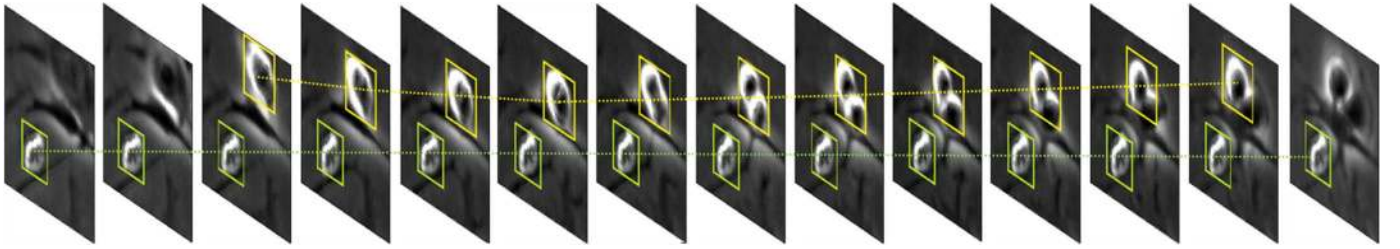


Fig. 2. An example of candidate patch sequences located in consecutive phase-contrast microscopy images. After candidate patch extraction based on brightness, spatially overlapped patches in consecutive frames are linked into a candidate patch sequence.

the result is thresholded with th , producing a binary image. For each connected component in the binary image, a bounding box is obtained and overlapping bounding boxes are then combined into one bounding box that encloses all the overlapping boxes. We then extract fixed-sized $s \times s$ rectangular patches from the image at the center position of each bounding box. The parameters d and s are closely related to the average cell size¹. In our experiments, we set parameters d , th , and s to be 10, 1000, and 50 using a typical cross validation scheme.

Following this, candidate patch sequences are constructed by associating spatially overlapped patches in consecutive frames as shown in Fig. 2. In the case that a patch in frame t can be linked with two or more patches in the following frame $t + 1$, the patch in frame t is associated with the one of the patches in frame $t + 1$, whereas each of the other patches in frame $t + 1$ initiates a new candidate patch sequence. In our experiments, multiple possible associations seldom occur except right after a birth event; a patch containing two new born cells that stick to each other can be linked with two patches each of which contains one of the daughter cells. In such a case, a birth event appears in the patch sequence before multiple possible associations; thus, association with either of the possibilities does not affect birth event detection performance.

B. Visual Feature Extraction

Unique scale gradient histogram features are extracted from each patch in candidate patch sequences. The unique scale gra-

¹Our experiments show that the best performance is achieved when parameters d and s are set to be approximately half and twice of the length of the square enclosing the average size of mitotic cells.

dient histogram features reflect the characteristics of phase-contrast microscopy images.

- *Relativity*: The pixel values in phase contrast microscopy images are influenced by many factors, such as artifacts and neighboring cells. Thus, relative features which reflect the change of pixel values, e.g., intensity gradient, are more reliable than absolute features.
- *Unique scale*: The variation of cell scales is minimal since the distances between the microscope lens and cells are almost uniform. Although cell sizes can differ within a certain range even during mitosis, the variation is limited and can thus be statistically modeled by training samples.
- *Rotation invariance*: Cells in a given field show various orientations. Regardless of the orientations, rotation invariance generally allows similar features to be extracted from cells with similar appearances.

The process computing unique scale gradient histograms follows scale invariance feature transform (SIFT) [19]. After dividing each patch into 4×4 subregions, we accumulate gradient magnitudes weighted by a Gaussian function into 4 bins along the orientations at each subregion as shown in Fig. 3. After $4 \times 4 \times 4 = 64$ features are computed for each patch, L2 normalization is applied to the feature vectors.

To achieve rotation invariance, each training candidate patch sequence is duplicated by rotating all patches in the sequence along several different orientations. This scheme results in performance improvement when training samples are insufficient. In our experiments, we applied three different orientations: 90° , 180° , and 270° as shown in Fig. 3. Other rotation schemes are not as effective as this simple duplication scheme because the

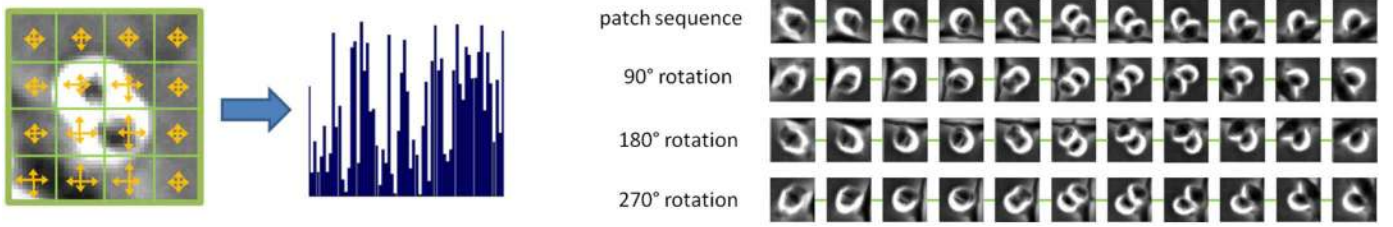


Fig. 3. Schematic of unique scale gradient histogram computation. At each of 4×4 subregions in a patch, a gradient histogram with four bins along the orientations are computed. To achieve rotation invariance, each candidate patch sequence in the training set is duplicated by rotating all patches in the sequence along three different orientations: 90° , 180° , and 270° .

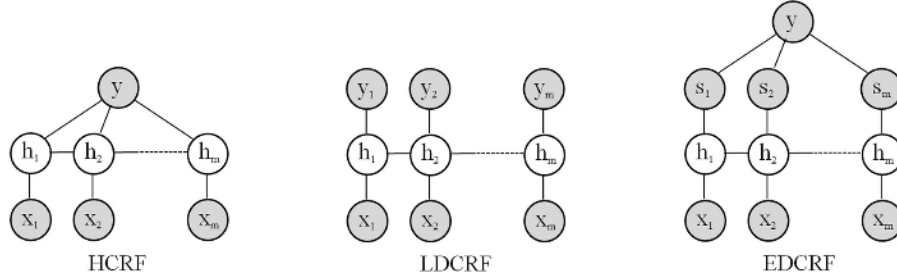


Fig. 4. Graphical representations of two previous models (HCRF and LDCRF) and our EDCRF model. x_i and h_i represent the i th observation (the i th patch in a given candidate sequence in our work) and the hidden state assigned on x_i , respectively. y represents a class label; one class label is assigned on the entire sequence in HCRF, while a class label is assigned for each of observations in LDCRF. In EDCRF, y is the label indicating the timing of the birth event and s_i is the i th sub-label determined by y . Gray circles denote observed variables for training. For testing, y and s are not observed.

major axis is not reliably found in phase-contrast microscopy images [19] or the relative spatial information is generally lost [20].

III. MITOSIS DECISION USING EVENT DETECTION CONDITIONAL RANDOM FIELD

After candidate patch sequence construction and visual feature extraction, the problem reduces to determining whether each candidate contains a birth event and which frame the birth event is located in. For these two decision tasks, we present event detection conditional random field (EDCRF), a probabilistic model for birth event detection and localization. After introducing related probabilistic models, we formulate the EDCRF model and describe its learning and inference process.

A. Related Models

We review two previous probabilistic models: hidden conditional random fields (HCRFs) [16] and latent-dynamic conditional random fields (LDCRFs) [18]. The graphical representations of these two models are shown in Fig. 4.

HCRF was devised to analyze temporal sequences, such as vision and speech applications [21], [22]. Particularly, it has been applied to gesture recognition tasks and demonstrates its superiority to hidden Markov models (HMM) and conditional random fields (CRF) [23]. HCRF also excels in mitosis occurrence detection [15]. However, since HCRF handles sequences on each of which only one label is imposed, it cannot capture the timing of particular events, such as birth events, in candidate patch sequences.

LDCRF was proposed to additionally capture extrinsic class dynamics based on the idea of HCRF [18]. Hidden variables in LDCRF not only model substructure of a class sequence, but also learn dynamics between class labels; thus, the model can

be directly applied to unsegmented sequences. Efficient training and inference schemes can be achieved by constraining each class label to have a disjoint set of associated hidden states. LDCRF has the potential to be modified into a model for event detection in that an event can be expressed as a class dynamic change when different labels are assigned on observations before and after the event.

B. EDCRF Formulation

EDCRF has one more variable containing event timing on the top of LDCRF as shown in Fig. 4. Given the event timing information, the sub-label variables in the subsequent layer are set to show an explicit label transition between before and after the event, which implies dynamic changes of observations. The formulation of EDCRF is as follows.

Suppose that n candidate patch sequence and label pairs $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$ are given. Each label y_i is defined as

$$y_i = \begin{cases} p, & \text{if the } p\text{th patch of } \mathbf{x}_i \text{ contains a birth event} \\ 0, & \text{if there exists no birth event in } \mathbf{x}_i \end{cases}. \quad (1)$$

Each sequence $\mathbf{x} = (x_1, x_2, \dots, x_m)$ consists of m candidate patches where x_j denotes the j th patch (m can be varied for different sequences). We assume hidden variables $\mathbf{h} = (h_1, h_2, \dots, h_m)$ and sub-labels $\mathbf{s} = (s_1, s_2, \dots, s_m)$ where h_j and s_j correspond to x_j . When a sequence label y is given, the sub-labels s_1, s_2, \dots, s_m are defined as

$$s_j = \begin{cases} N, & \text{if } y = 0 \\ B, & \text{if } y > 0 \text{ and } j < y \\ A, & \text{if } y > 0 \text{ and } j \geq y \end{cases} \quad (2)$$

where label N , B , and A represent *no event*, *before the event*, and *after the event (including the event)*, respectively. In other words,

if there exists a birth event in a given candidate sequence, the sub-labels before the event are set to be B and the other sub-labels set to be A . Otherwise, all the sub-labels are set to be N .

Under these definitions, we define a latent conditional model for each sequence

$$P(y|\mathbf{x}, \theta) = P(\mathbf{s}|\mathbf{x}, \theta) = \sum_{\mathbf{h}} P(\mathbf{s}|\mathbf{h}, \mathbf{x}, \theta) P(\mathbf{h}|\mathbf{x}, \theta) \quad (3)$$

where θ is a set of parameters of the model.

In order to make the modeling efficient, we adopt the same scheme as LDCRF [18], which restricts that each sub-class label s is associated only with hidden states in a disjoint set \mathcal{H}_s . Then

$$P(\mathbf{s}|\mathbf{h}, \mathbf{x}, \theta) = \begin{cases} 1, & \text{if } \forall h_j \in \mathcal{H}_{s_j} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

The proposed model is thus simplified as

$$P(y|\mathbf{x}, \theta) = \sum_{\mathbf{h}: \forall h_j \in \mathcal{H}_{s_j}} P(\mathbf{h}|\mathbf{x}, \theta). \quad (5)$$

We define $P(\mathbf{h}|\mathbf{x}, \theta)$ using the typical conditional random field (CRF) formulation

$$P(\mathbf{h}|\mathbf{x}, \theta) = \frac{1}{Z} \exp \left(\sum_{j=1}^m f^{(s)}(h_j, \mathbf{x}, j) \cdot \theta^{(s)}(h_j) + \sum_{j=2}^m f^{(t)}(h_{j-1}, h_j, \mathbf{x}, j) \cdot \theta^{(t)}(h_{j-1}, h_j) \right) \quad (6)$$

where Z is a partition function. $f^{(s)}(h_j, \mathbf{x}, j)$ and $f^{(t)}(h_{j-1}, h_j, \mathbf{x}, j)$ are a state and a transition function, respectively. $\theta^{(s)}$ and $\theta^{(t)}$ are the parameters of state and transition functions, respectively. Also, $\theta = \{\theta^{(s)}, \theta^{(t)}\}$.

We define state functions as

$$f^{(s)}(h_j, \mathbf{x}, j) = \phi(x_j) \quad (7)$$

where $\phi(x_j)$ is a visual feature vector of x_j . The inner product of $\phi(x_j) \cdot \theta^{(s)}(h_j)$ can be interpreted as a compatibility measure between observation x_j and hidden state h_j [16]. We define transition functions as

$$f^{(t)}(h_{j-1}, h_j, \mathbf{x}, j) = \begin{cases} 1, & \text{if } (h_{j-1}, h_j) \in \mathcal{U} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

where $\mathcal{U} = \mathcal{H}_N \times \mathcal{H}_N \cup \mathcal{H}_B \times \mathcal{H}_B \cup \mathcal{H}_B \times \mathcal{H}_A \cup \mathcal{H}_A \times \mathcal{H}_A$. We restrict sub-level transitions among (N, N) , (B, B) , (B, A) , and (A, A) , which respectively represent *no event*, *before the event*, *during the event*, and *after the event*. Aside from these four transitions, there exists no other transition in our setting.

C. Learning Model Parameters

For learning parameters, we maximize the following regularized log-likelihood function as conventionally [24], [25]

$$L(\theta) = \sum_{i=1}^n \log P(y_i|\mathbf{x}_i, \theta) - \frac{1}{2\sigma^2} \|\theta\|^2 \quad (9)$$

where σ is the variance of a Gaussian prior.

This optimization problem can be solved by gradient ascent methods. The derivative of $L(\theta)$ with respect to $\theta_k^{(s)}(h')$, which is the k th element of $\theta^{(s)}(h')$, is computed as

$$\begin{aligned} & \frac{\partial L(\theta)}{\partial \theta_k^{(s)}(h')} \\ &= \sum_{i=1}^n \sum_{j=1}^m P(h_j = h' | y_i, \mathbf{x}_i, \theta) f_k^{(s)}(h', \mathbf{x}_i, j) \\ & \quad - \sum_{i=1}^n \sum_{y'=0}^m \sum_{j=1}^m P(h_j = h', y' | \mathbf{x}_i, \theta) \\ & \quad \times f_k^{(s)}(h', \mathbf{x}_i, j) - \frac{\theta_k^{(s)}(h')}{\sigma^2} \end{aligned} \quad (10)$$

where $f_k^{(s)}(h, \mathbf{x}_i, j)$ is the k th element of $f^{(s)}(h, \mathbf{x}_i, j)$. $P(h_j = h' | y_i, \mathbf{x}_i, \theta)$ can be computed by belief propagation [26] in $O(m)$ [18].

Similarly, the derivative of $L(\theta)$ with respect to $\theta^{(t)}(h', h'')$ is computed as

$$\begin{aligned} & \frac{\partial L(\theta)}{\partial \theta^{(t)}(h', h'')} \\ &= \sum_{i=1}^n \sum_{j=2}^m P(h_{j-1} = h', h_j = h'' | y_i, \mathbf{x}_i, \theta) \\ & \quad \times f^{(t)}(h', h'', \mathbf{x}_i, j) \\ & \quad - \sum_{i=1}^n \sum_{y'=0}^m \sum_{j=2}^m P(h_{j-1} = h', h_j = h'', y' | \mathbf{x}_i, \theta) \\ & \quad \times f^{(t)}(h', h'', \mathbf{x}_i, j) \\ & \quad - \frac{\theta^{(t)}(h', h'')}{\sigma^2}. \end{aligned} \quad (11)$$

For valid transition $(h_{j-1} = h', h_j = h'')$, since $f^{(t)} = 1$, the derivative can be simplified as

$$\begin{aligned} & \frac{\partial L(\theta)}{\partial \theta^{(t)}(h', h'')} \\ &= \sum_{i=1}^n \sum_{j=2}^m P(h_{j-1} = h', h_j = h'' | y_i, \mathbf{x}_i, \theta) \\ & \quad - \sum_{i=1}^n \sum_{y'=0}^m \sum_{j=2}^m P(h_{j-1} = h', h_j = h'', y' | \mathbf{x}_i, \theta) \\ & \quad - \frac{\theta^{(t)}(h', h'')}{\sigma^2}. \end{aligned} \quad (12)$$

$P(h_{j-1} = h', h_j = h'' | y_i, \mathbf{x}_i, \theta)$ can also be efficiently computed by belief propagation.

D. Inferences

For testing of a new sequence \mathbf{x} , we first compute the probabilities of our conditional model with all possible y and the optimal parameter θ^* obtained in the training step

$P(y = 0 | \mathbf{x}, \theta^*)$ can be computed as

$$\begin{aligned} & P(y = 0 | \mathbf{x}, \theta^*) \\ &= P(s_1 = N, \dots, s_m = N | \mathbf{x}, \theta^*) \\ &= P(s_1 = N | \mathbf{x}, \theta^*) = \sum_{h_1 \in H_N} P(h_1 | \mathbf{x}, \theta^*) \end{aligned} \quad (13)$$

TABLE 1
COMPARISON OF MITOSIS DETECTION PERFORMANCE BETWEEN EDCRF AND HCRF MODELS WHEN THE TIMING OF BIRTH EVENTS IS NOT CONSIDERED. COMPARISON IS CONDUCTED IN TERMS OF PRECISION, RECALL, F-MEASURE, AND THE AUC OF THE PR-CURVE ON FOUR C3H10T1/2 AND ONE C2C12 STEM CELL POPULATIONS. EDCRF OUTPERFORMS HCRF IN TERMS OF ALL THE METRICS

Model	C3H10T1/2				C2C12			
	Precision	Recall	F-measure	AUC	Precision	Recall	F-measure	AUC
EDCRF	0.913 \pm 0.040	0.870 \pm 0.048	0.889 \pm 0.007	0.958 \pm 0.006	0.950	0.893	0.921	0.930
HCRF	0.890 \pm 0.019	0.863 \pm 0.033	0.876 \pm 0.010	0.952 \pm 0.005	0.924	0.875	0.899	0.866

because $s_1 = N$ leads to $s_2, \dots, s_m = N$ under our restricted transition rule. Similarly

$$\begin{aligned}
 P(y = 1|\mathbf{x}, \theta^*) &= P(s_1 = A, \dots, s_m = A|\mathbf{x}, \theta^*) \\
 &= P(s_1 = A|\mathbf{x}, \theta^*) = \sum_{h_1 \in H_A} P(h_1|\mathbf{x}, \theta^*). \quad (14)
 \end{aligned}$$

The other conditional probabilities can be computed as

$$\begin{aligned}
 P(y = j|\mathbf{x}, \theta^*) &= P(s_1 = B, \dots, s_{j-1} = B, s_j = A, \dots, s_m = A|\mathbf{x}, \theta^*) \\
 &= P(s_{j-1} = B, s_j = A|\mathbf{x}, \theta^*) \\
 &= P(s_{j-1} = B|\mathbf{x}, \theta^*) - P(s_{j-1} = B, s_j = B|\mathbf{x}, \theta^*) \\
 &= P(s_{j-1} = B|\mathbf{x}, \theta^*) - P(s_j = B|\mathbf{x}, \theta^*) \\
 &= \sum_{h_{j-1} \in H_B} P(h_{j-1}|\mathbf{x}, \theta^*) - \sum_{h_j \in H_B} P(h_j|\mathbf{x}, \theta^*) \\
 &\quad \text{for } j = 2, \dots, m. \quad (15)
 \end{aligned}$$

For mitosis occurrence decision on each candidate sequence, we compare $P(y = 0|\mathbf{x}, \theta^*)$ and $1 - P(y = 0|\mathbf{x}, \theta^*)$. If the former is greater, EDCRF determines that there is no mitotic event in the given sequence. Otherwise, the temporal localization of the birth event follows by comparing $P(y = 1|\mathbf{x}, \theta^*)$, \dots , and $P(y = m|\mathbf{x}, \theta^*)$. More formally,

$$y^* = \begin{cases} 0, & \text{if } P(y = 0|\mathbf{x}, \theta^*) > 0.5 \\ \arg \max_{y=1, \dots, m} P(y|\mathbf{x}, \theta^*), & \text{otherwise} \end{cases} \quad (16)$$

IV. EXPERIMENTAL SETUP

A. Data and Ground Truth

Multipotent C3H10T1/2 mesenchymal stem cells (ATTC, Manassas, VA) serve as a model for the adult human mesenchymal stem cell and were grown in Dulbecco's Modified Eagle's Media (DMEM; Invitrogen, Carlsbad, CA), 10% fetal bovine serum (Invitrogen, Carlsbad, CA) and 1% penicillin-streptomycin (PS; Invitrogen, Carlsbad, CA). C2C12 myoblastic stem cells (ATTC, Manassas, VA) have the capacity to differentiate into osteoblasts and myocytes and were grown in DMEM, 10% bovine serum (Invitrogen, Carlsbad, CA) and 1% PS. All cells were kept at 37 °C, 5% CO₂ in a humidified incubator.

Phase-contrast microscopy images of the two types of stem cell populations (C3H10T1/2 and C2C12) were generated as follows. During the growth of stem cells, microscopy cell images were acquired every 5 min using a Zeiss Axiovert

T135V microscope (Carl Zeiss Microimaging, Thornwood, NY) equipped with a 5X, 0.15 N.A. phase-contrast objective, a custom-stage incubator, and the InVitro software (Media Cybernetics Inc., Bethesda, MD). Each of the images contains 1392 \times 1040 pixels with a resolution of 1.3 μ m/pixel. C3H10T1/2 and C2C12 microscopy image sequences consist of 1436 and 1013 images, respectively.

After acquiring the image sequences, manual annotation of birth events was performed on one C2C12 and five C3H10T1/2 image sequences. For each birth event, the center of the boundary between two daughter cells was marked when the boundary is clearly observed. Since C2C12 myoblasts were cultured to a much higher level of confluence than C3H10T1/2 mesenchymal stem cells in our data, each of the C3H10T1/2 sequences contain 41–128 mitotic events, while the C2C12 sequence contains 673 mitotic events.

B. Experiments

For mitosis detection without considering the timing of birth events, we compare EDCRF with HCRF [16], which was previously used for mitosis occurrence detection. HCRF is known to outperform hidden Markov models (HMM) and conditional random fields (CRF) [15], [23].

To the best of our knowledge, there exists no probabilistic model which has been used for automated temporal localization of birth events given candidate patch sequences; in this respect, our EDCRF model is original. In order to compare EDCRF with possible alternatives, we additionally use either support vector machines (SVM) or conditional random fields (CRF) for the temporal localization of the birth event after the identification of mitosis occurrence using HCRF. In these alternative approaches, the identification of mitosis occurrence and the temporal localization of birth event are sequentially performed.

In the model incorporating HCRF and SVM (HCRF + SVM), a version of SVM that outputs probabilities is applied to the sequences determined by HCRF to have a mitosis occurrence. More specifically, for training, candidate patches containing manual annotation of birth events are used as positive samples and all patches in the candidate patch sequences that do not contain mitosis as negative samples. For testing, the SVM produces the probabilities that each patch contains a birth event; the patch with the highest probability in the sequence is then decided to contain a birth event.

In the model combining HCRF and CRF (HCRF + CRF), the same labeling scheme as EDCRF is applied to temporally localize birth events. More specifically, for training, one label is assigned to the label variables before the birth event and another label is assigned after the event. If there exists no birth event, a third label is assigned to all the label variables. For testing,

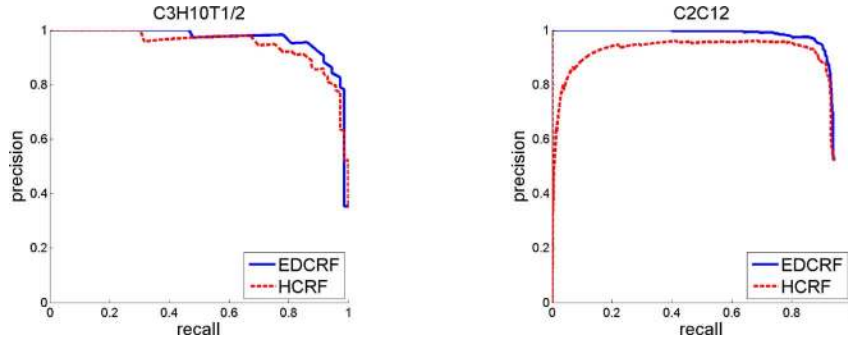


Fig. 5. The PR-curves of mitosis detection on the first C3H10T1/2 and the C2C12 image sequences when the timing of birth events is not considered. In terms of AUC, the EDCRF model outperforms the HCRF model in both cell types.

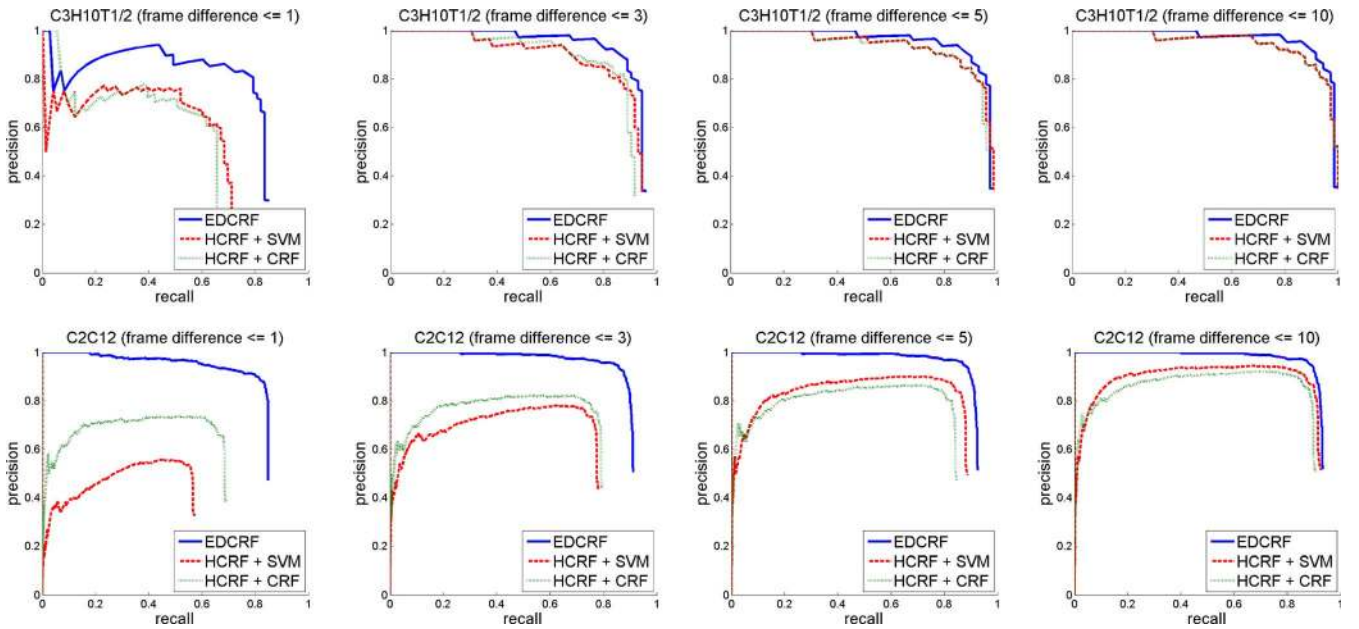


Fig. 6. The PR-curves of EDCRF and two alternative models (HCRF + SVM and HCRF + CRF) showing mitosis detection accuracy on the first C3H10T1/2 (top) and the C2C12 (bottom) image sequences. Detection results are considered true positive when the timing error of the birth event is not greater than a given threshold (one of 1, 3, 5, and 10). In terms of AUC, our approach consistently outperforms the alternative approaches regardless of the threshold. The superiority of our approach is more evident when the threshold is small, i.e., the evaluation on birth event timing is strict.

the same inference scheme as the EDCRF model is utilized to determine a birth event's temporal location.

For C3H10T1/2 image data, one sequence is used for training and the other four sequences for testing whereas for C2C12 image data, half of all mitotic cells are used for training in turn, and the other half are used for testing.

C. Evaluation

After constructing candidate patch sequences, the number of undetected mitosis is counted. Each mitotic event, specifically the birth event, is expected to be captured by one of the candidate patch sequences. However, if there are birth events which are not contained in any candidate patch sequence, such cases are considered undetected. If one candidate patch sequence contains more than one birth event, all of the birth events except the first one are considered undetected as well. All of the undetected mitosis are counted as false negatives when precision and recall are computed.

After applying the probabilistic models, we first evaluated the identification of mitosis occurrence in terms of precision and re-

call without examining the timing of birth events. In this case, true positive is defined as the case that a candidate patch sequence containing a birth event is correctly identified no matter how great the timing error of the birth event is. If one birth event is located in more than one candidate patches, the birth event is considered to exclusively belong to the patch whose center is the closest to the birth event among the candidate patches.

We then evaluate the identification of mitosis occurrence with the constraint of birth event timing. In this case, among the aforementioned true positive cases, only the cases in which the timing error of the birth event is equal to or less than a certain threshold are considered true positive. In other words, although a candidate sequence including an actual birth event is correctly identified, if the timing error of the birth event is greater than the threshold, the birth event is considered undetected as well as the detection is regarded false. More specifically, the timing error is measured as the frame difference between the patch containing the ground truth and the patch containing the detected birth event in the sequence. We use four different thresholds (1, 3, 5, and 10) and report precision and recall for each

TABLE II

MITOSIS DETECTION PERFORMANCE COMPARISON WHEN THE TIMING OF BIRTH EVENTS IS EXAMINED IN ADDITION TO MITOSIS OCCURRENCE. OUR APPROACH IS COMPARED WITH THE ALTERNATIVE APPROACHES BASED ON HCRF (HCRF + SVM AND HCRF + CRF) IN TERMS OF PRECISION, RECALL, F-MEASURE, AND THE AUC OF THE PR-CURVE ON C3H10T1/2 AND C2C12 STEM CELL POPULATIONS. DETECTION RESULTS ARE CONSIDERED TRUE POSITIVE WHEN THE TIMING ERROR OF THE BIRTH EVENT IS NOT GREATER THAN A GIVEN THRESHOLD (ONE OF 1, 3, 5, AND 10). WHEN A THRESHOLD IS SMALL (THE EVALUATION IS STRICT), OUR APPROACH MORE CLEARLY OUTPERFORMS THE ALTERNATIVES

Threshold of frame difference	EDCRF							
	C3H10T1/2				C2C12			
	Precision	Recall	F-measure	AUC	Precision	Recall	F-measure	AUC
1	0.740 \pm 0.067	0.703 \pm 0.035	0.720 \pm 0.041	0.660 \pm 0.047	0.880	0.828	0.853	0.819
3	0.863 \pm 0.044	0.822 \pm 0.043	0.840 \pm 0.015	0.882 \pm 0.026	0.925	0.870	0.897	0.895
5	0.899 \pm 0.037	0.857 \pm 0.052	0.876 \pm 0.014	0.934 \pm 0.017	0.940	0.884	0.911	0.913
10	0.910 \pm 0.042	0.867 \pm 0.047	0.886 \pm 0.009	0.952 \pm 0.013	0.947	0.890	0.918	0.923

Threshold of frame difference	HCRF + SVM							
	C3H10T1/2				C2C12			
	Precision	Recall	F-measure	AUC	Precision	Recall	F-measure	AUC
1	0.604 \pm 0.052	0.585 \pm 0.051	0.594 \pm 0.049	0.466 \pm 0.055	0.550	0.520	0.535	0.270
3	0.795 \pm 0.045	0.773 \pm 0.064	0.783 \pm 0.052	0.785 \pm 0.090	0.765	0.724	0.744	0.552
5	0.872 \pm 0.010	0.847 \pm 0.043	0.859 \pm 0.019	0.920 \pm 0.022	0.873	0.826	0.849	0.745
10	0.887 \pm 0.019	0.860 \pm 0.033	0.873 \pm 0.010	0.946 \pm 0.007	0.909	0.861	0.884	0.828

Threshold of frame difference	HCRF + CRF							
	C3H10T1/2				C2C12			
	Precision	Recall	F-measure	AUC	Precision	Recall	F-measure	AUC
1	0.583 \pm 0.044	0.565 \pm 0.036	0.574 \pm 0.037	0.463 \pm 0.031	0.687	0.650	0.668	0.473
3	0.771 \pm 0.055	0.749 \pm 0.071	0.759 \pm 0.060	0.744 \pm 0.099	0.788	0.746	0.767	0.612
5	0.853 \pm 0.036	0.827 \pm 0.039	0.839 \pm 0.028	0.890 \pm 0.043	0.838	0.793	0.815	0.685
10	0.884 \pm 0.020	0.857 \pm 0.034	0.870 \pm 0.012	0.943 \pm 0.018	0.892	0.845	0.868	0.787

case. The smaller a threshold is, the stricter the evaluation is. To compare the detection results, we also compute F-measure, which is the harmonic mean of precision and recall, and the area under the curve (AUC) of the precision-recall curve (PR-curve). PR-curves are obtained by varying the decision probability in (16).

V. RESULTS AND DISCUSSIONS

During candidate patch sequence extraction, one birth event is not captured in four C3H10T1/2 sequences which are used for testing. There is no case that one candidate patch sequence contains more than one birth event in the C3H10T1/2 sequences. In the case of the C2C12 sequence, one birth event is missed and 36 birth events are detected following another birth event in the same sequence; as a result, a total of 37 false negative cases are reported before the decision tasks. Multiple mitosis in a candidate patch sequence occur due to the adhesion of mitotic cells at high confluency present in the C2C12 sequence. Under such circumstances, it is difficult to identify attached cells as separate entities.

As shown in Table I, the proposed mitosis detection method achieves 0.913/0.870 and 0.950/0.893 in terms of precision/recall on C3H10T1/2 and C2C12 stem cell populations, respectively, when only mitosis occurrence is considered. In terms of F-measure and AUC, the accuracy on C2C12 is comparable to that of C3H10T1/2, although the C2C12 cell population is more challenging due to its higher level of confluence and deformability. Having more training samples for C2C12 might be the reason.

Compared with the HCRF model, the EDCRF model is superior in mitosis occurrence detection in terms of precision, recall, F-measure, and the AUC of the PR-curve as shown in Table I and Fig. 5. A Student's paired t-test on the F-measures shows

that the performance improvement is statistically significant at the significance level 0.01 ($p = 0.0008$). These results indicate that the information of birth event timing is actually helpful for identifying the occurrence of mitosis. HCRF cannot utilize such additional information due to its limited expression power. On the other hand, EDCRF simultaneously models both mitosis occurrence and birth event timing, resulting in higher discrimination power than HCRF in mitosis occurrence identification.

When we additionally consider the timing errors of birth events and threshold them to obtain true positive cases, the superiority of EDCRF is more obvious. Compared to the alternative models, HCRF + SVM and HCRF + CRF, EDCRF consistently outperforms them in terms of precision, recall, F-measure, and AUC regardless of the cell type and the threshold for the timing error of birth events as shown in Table II. Student's paired t-tests on the F-measures show that the performance improvements are statistically significant at the significance level 0.01 regardless of the threshold of the timing error (EDCRF versus HCRF + SVM : $p = 0.0024, 0.0038, 0.0092$, and 0.0028 ; EDCRF versus HCRF + CRF : $p = 0.0001, 0.0013, 0.0055$, and 0.0037 for the threshold = 1, 3, 5, and 10, respectively). When a smaller threshold corresponding to a stricter evaluation of temporal localization is applied, our approach significantly outperforms the alternative approaches as seen in Fig. 6. In the alternative models, the identification of mitosis occurrence and the temporal localization of birth events are separately performed so the localization step may not be meaningful if mitosis occurrence is incorrectly identified. In this sense, the preceding mitosis occurrence decision may undermine the full potential of the localization step that follows. It is worth mentioning that recalls do not reach one in the PR-curves because some of existing mitosis are not detected regardless of the decision probability of mitosis occurrence. The undetected

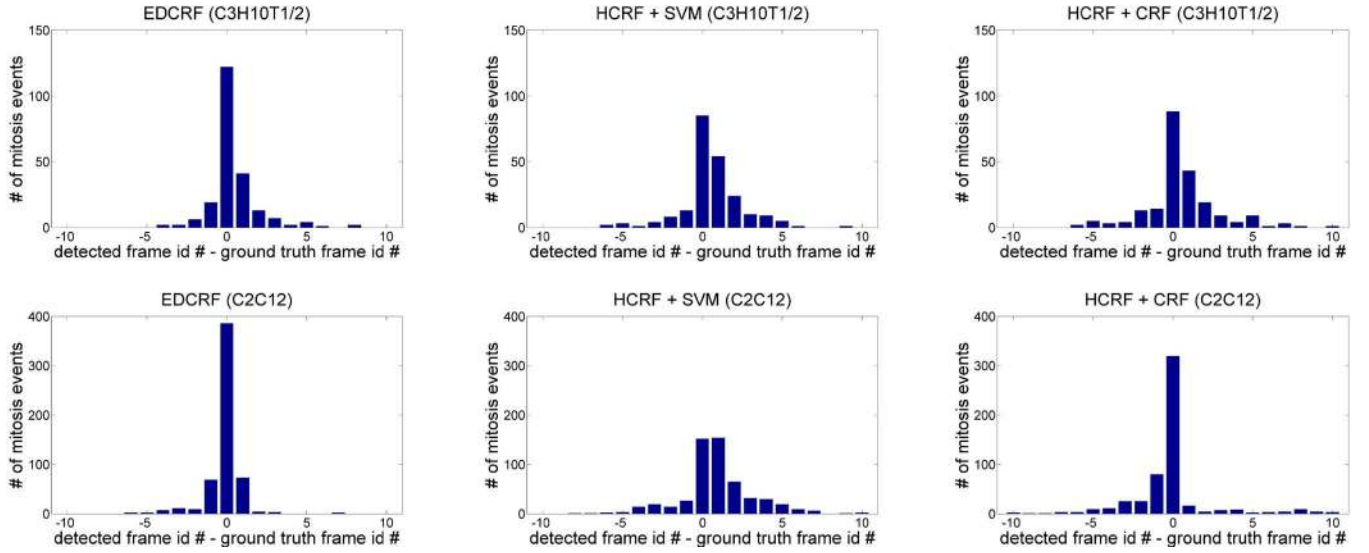


Fig. 7. Temporal localization precision of birth events of three approaches: EDCRF (left), HCRF + SVM (middle), and HCRF + CRF (right) on two cell populations: C3H10T1/2 (top) and C2C12 (bottom). The histograms show the frequency distribution of birth event timing errors when at most 10 frame error is allowed. Timing error is measured as the frame difference between the patches containing the ground truth and the detected result of the birth event among true positive samples. The temporal localization of EDCRF is more accurate than the alternatives.

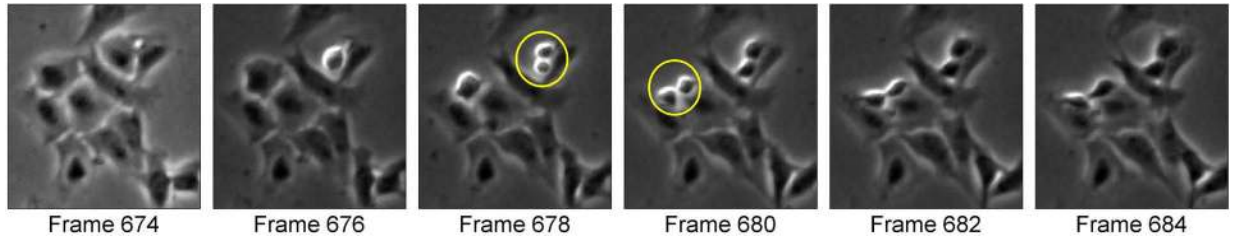


Fig. 8. Sample images illustrating examples of mitosis detection on the C2C12 microscopy image sequence. Each of the yellow circles in frames 678 and 680 surrounds a detected birth event. The EDCRF model temporally localizes birth events when there is a distinct boundary between daughter cells.

TABLE III

AVERAGE AND STANDARD DEVIATION OF THE TIMING ERROR OF BIRTH EVENTS IN TERMS OF (ABSOLUTE) FRAME DIFFERENCE. THE ERROR OF EDCRF IS SMALLER THAN THE ALTERNATIVE MODELS. THE IMPROVEMENT IN PERFORMANCE IS STATISTICALLY SIGNIFICANT

Cell type	EDCRF	HCRF+SVM	HCRF+CRF
C3H10T1/2	0.828 \pm 1.334	1.273 \pm 1.528	1.438 \pm 1.826
C2C12	0.482 \pm 0.957	1.685 \pm 1.769	1.189 \pm 2.120

mitosis occur due to either imperfect extraction of candidate patch sequences or inaccurate temporal localization of birth events.

Fig. 7 shows the distribution of the frame differences between the human and computer annotations of birth events among true positive samples on C3H10T1/2 and C2C12 sequences when at most 10 frame difference is allowed in temporal localization of birth events. The averages of the (absolute) frame differences using EDCRF are smaller than using the alternative models as shown in Table III. EDCRF achieves statistically significant improvements in the temporal localization of birth events; when applying Student's *t*-tests, we obtain *p*-values less than 0.001 for all of the four cases: comparison with HCRF + SVM on C3H10T1/2 ($p = 0.0006$), HCRF + CRF on C3H10T1/2 ($p < 0.0001$), HCRF + SVM on C2C12 ($p < 0.0001$), and HCRF +

CRF on C2C12 ($p < 0.0001$). In addition to the lesser performance of HCRF in the identification of mitosis occurrence, SVM and CRF also fall short of the capability of EDCRF in the temporal localization of birth events. SVM is not capable of modeling temporal dynamic and CRF does not capture the hidden state structures in candidate patch sequences.

Using our design, the overall process of mitosis detection is more computationally efficient compared to the previous work [15]. By removing the time-consuming preconditioning [17], our approach can process a test image of 1392×1040 pixels in less than 5 s, while the previous method spends more than 5 min on the preconditioning step alone when using a computer with a dual core 2-GHz processor and 2-GB memory. This computational improvement enables real-time analysis of microscopy images periodically taken even with a short time interval. The previously used preconditioning scheme [17] is not required because cells show distinctive characteristics during mitosis. Mitotic cells can be recognized without applying the complex method devised for nonmitotic cells rather than mitotic cells.

Fig. 8 demonstrates two examples of birth events in the C2C12 myoblastic stem cell population automatically detected by our approach. Our approach shows good performance on the

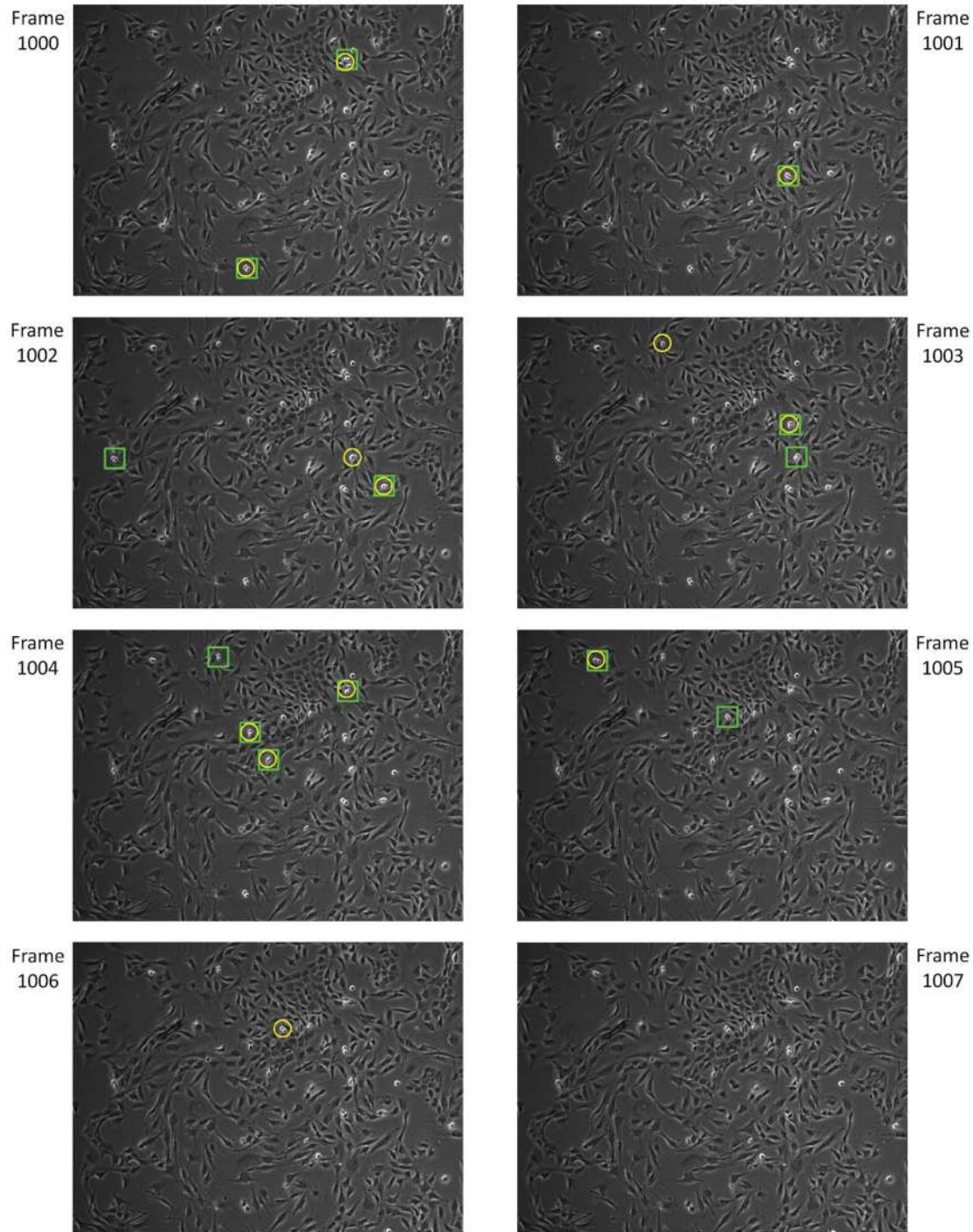


Fig. 9. Sample images illustrating examples of mitosis detection on the C2C12 microscopy image sequence from frames 1000 through 1007. Each of the yellow circles surrounds a detected birth event produced by our approach, whereas each of green squares encloses a ground truth birth event annotated by a human. During these eight frames, twelve birth events are correctly detected within one frame error and one birth event is missed in frame 1002.

C2C12 stem population with higher confluency (approximately 80%–90%) as shown in Fig. 9.

VI. CONCLUSION AND FUTURE WORK

We have proposed an effective approach for mitosis detection, specifically birth event detection and temporal localization. Mitosis detection accuracy and speed are considerably improved compared to previous work by developing a probabilistic model for event detection and bypassing the use of time consuming

modules. Experimental results on two types of stem cell populations validate the efficacy of our approach.

The mitosis detection algorithm described here will facilitate the quantification of biological metrics for cell proliferation. Such quantification will be useful in the study and characterization of biological processes and may have applications in high throughput screens that are reliant on cell proliferation measurements including identification of potent anti-mitotic drugs for chemotherapy as well as determination of drug sensitivity

in normal and diseased cells. In addition, we plan to develop a real-time cell tracking system that incorporates the functionality of our mitosis detection approach. The system would facilitate online monitoring and adaptive control of stem cell expansions by helping to make online decisions about when to subculture the stem cells as well as alerting the users of abnormalities of cell behaviors.

ACKNOWLEDGMENT

The authors would like to thank P. Campbell, L. Weiss, and Z. Yin for providing helpful comments on the manuscript.

REFERENCES

- [1] H. Quastler and F. G. Sherman, "Cell population kinetics in the intestinal epithelium of the mouse," *Exp. Cell Res.*, vol. 17, no. 3, pp. 420–438, 1959.
- [2] "Assays for cell proliferation studies," *Genet. Eng. Biotechnol.* N. vol. 26, no. 6, 2006 [Online]. Available: <http://www.genengnews.com/gen-articles/assays-for-cell-proliferation-studies/1442/>
- [3] C. Zimmer, E. Labruière, V. Meas-Yedid, N. Guillén, and J.-C. Olivo-Marin, "Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: A tool for cell-based drug testing," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1212–1221, Oct. 2002.
- [4] A. J. Hand, T. Sun, D. C. Barber, D. R. Hose, and S. Macneil, "Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration," *J. Microsc.*, vol. 234, no. 1, pp. 62–79, 2008.
- [5] K. Li, E. D. Miller, M. Chen, T. Kanade, L. E. Weiss, and P. G. Campbell, "Computer vision tracking of stemness," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2008, pp. 847–850.
- [6] D. House, M. L. Walker, Z. Wu, J. Y. Wong, and M. Betke, "Tracking of cell populations to understand their spatio-temporal behavior in response to physical stimuli," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshop Math. Methods Biomed. Image Anal.*, 2009, pp. 186–193.
- [7] F. Yang, M. A. Mackey, F. Ianzini, G. Gallardo, and M. Sonka, "Cell segmentation, tracking, and mitosis detection using temporal context," in *Proc. Int. Conf. Med. Image Computing Computer Assist. Intervent.*, 2005, pp. 302–309.
- [8] J. Huth, M. Buchholz, J. M. Kraus, M. Schmucker, G. von Wichert, D. Krndija, T. Seufferlein, T. M. Gress, and H. A. Kestler, "Significantly improved precision of cell migration analysis in time-lapse video microscopy through use of a fully automated tracking system," *BMC Cell Biol.*, vol. 11, no. 1, pp. 24–35, 2005.
- [9] O. Debeir, P. Van Ham, R. Kiss, and C. Decaestecker, "Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes," *IEEE Trans. Med. Imag.*, vol. 24, no. 6, pp. 697–711, Jun. 2005.
- [10] O. Al-Kofahi, R. J. Radke, S. K. Goderie, Q. Shen, S. Temple, and B. Roysam, "Automated cell lineage construction: A rapid method to analyze clonal development established with murine neural progenitor cells," *Cell Cycle*, vol. 5, no. 3, pp. 327–335, 2006.
- [11] D. Padfield, J. Rittscher, N. Thomas, and B. Roysam, "Spatio-temporal cell cycle phase analysis using level sets and fast marching methods," *Med. Image Anal.*, vol. 13, no. 1, pp. 143–155, 2009.
- [12] O. Debeir, V. Mégalizzi, N. Warzée, R. Kiss, and C. Decaestecker, "Videomicroscopic extraction of specific information on cell proliferation and migration *in vitro*," *Exp. Cell Res.*, vol. 314, no. 16, pp. 2985–2998, 2008.
- [13] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. Comput. Vis. Pattern Recognit.*, 2001, pp. 511–518.
- [14] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997.
- [15] A.-A. Liu, K. Li, and T. Kanade, "Mitosis sequence detection using hidden conditional random fields," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2010.
- [16] A. Quattoni, S. Wang, L. Morency, M. Collins, and T. Darrell, "Hidden conditional random fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1848–1853, Oct. 2007.
- [17] K. Li and T. Kanade, "Nonnegative mixed-norm preconditioning for microscopy image segmentation," in *Proc. Int. Conf. Inf. Process. Med. Imag.*, 2009, pp. 362–373.
- [18] L. P. Morency, A. Quattoni, and T. Darrell, "Latent-dynamic discriminative models for continuous gesture recognition," in *Proc. Comput. Vis. Pattern Recognit.*, 2007, pp. 1–8.
- [19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [20] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1265–1278, Aug. 2005.
- [21] A. Quattoni, M. Collins, and T. Darrell, "Conditional random fields for object recognition," in *Proc. Neural Inf. Process. Syst.*, 2004, pp. 1097–1104.
- [22] A. Gunawardana, M. Mahajan, A. Acero, and J. C. Platt, "Hidden conditional random fields for phone classification," *Interspeech*, pp. 1117–1120, 2005.
- [23] S. B. Wang, A. Quattoni, L.-P. Morency, D. Demirdjian, and T. Darrell, "Hidden conditional random fields for gesture recognition," in *Proc. Comput. Vis. Pattern Recognit.*, 2006, pp. 1521–1527.
- [24] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labelling sequence data," in *Proc. Int. Conf. Mach. Learn.*, 2001, pp. 282–289.
- [25] S. Kumar and M. Herbert, "Discriminative random fields: A framework for contextual interaction in classification," in *Proc. Int. Conf. Comput. Vis.*, 2003, pp. 1150–1157.
- [26] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann, 1988.