MITOSIS SEQUENCE DETECTION USING HIDDEN CONDITIONAL RANDOM FIELDS

A-A Liu[†], K. Li[‡], T. Kanade

Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213

ABSTRACT

We propose a fully-automated mitosis event detector using hidden conditional random fields for cell populations imaged with time-lapse phase contrast microscopy. The method consists of two stages that jointly optimize recall and precision. First, we apply model-based microscopy image preconditioning and volumetric segmentation to identify candidate spatiotemporal sub-regions in the input image sequence where mitosis potentially occurred. Then, we apply a learned hidden conditional random field classifier to classify each candidate sequence as mitosis or not. The proposed detection method achieved 95% precision and 85% recall in very challenging image sequences of multipolar-shaped C3H10T1/2 mesenchymal stem cells. The superiority of the method was further demonstrated by comparisons with conditional random field and support vector machine classifiers. Moreover, the proposed method does not depend on empirical parameters, *ad hoc* image processing, or cell tracking; and can be straightforwardly adapted to different cell types.

Index Terms— Mitosis, Hidden Conditional Random Field, Image Preconditioning, Phase Contrast Microscopy

1. INTRODUCTION

Measurement of the proliferative behaviors of cells *in vitro* is important to many biomedical applications ranging from basic biological research to advanced applications, such as drug discovery, stem cell manufacturing, and tissue engineering. Critical to such measurement is the accurate counting and localization of occurrences of *mitosis*, or cell division, in a cell culture. For short-period, small-scale studies, it is possible to manually identify incidents of mitosis because mitotic cells in culture tend to retract, round up, and exhibit intensified surrounding halos under phase contrast illumination. However, the need for extended-time observation and the proliferation of high-throughput imaging have made automated image analysis mandatory.

Automated mitosis detection methods in prior art can be categorized into tracking-based, tracking-free, and hybrid approaches. *Tracking-based* approaches [1] rely on cell tracking to determine individual cell trajectories, and then identify mitosis based on the temporal progression of cell features along their trajectories. The dependency on cell tracking is a severe burden because tracking *per se* is a challenging task. *Tracking-free* approaches alleviate this burden and can detect mitosis directly in an image sequence. One representative technique was proposed by Li *et al* [4], which applies a cascade classifier to classify volumetric sliding windows of an image sequence with 3D Haar-like features. Major drawbacks of this approach include the requirement of a large amount of training data and the lack of location specificity of detection. *Hybrid* approaches aim to construct a self-contained solution by leveraging the advantages of the previous two methods. These approaches typically consist of candidate sequence detection, sequence feature extraction, and classification as consecutive steps. To detect mitosis candidates, earlier methods [5] apply thresholding and morphological filtering to extract bright halos surrounding potentially mitotic cells in each image, and then group the extracted regions in successive images based on their spatial relationship. Subsequently, to identify mitosis, Eccles *et al* [5] employed a ring shape detector to locate the mother and two daughter cells; Gallardo *et al* [6] adopted a hidden Markov model to classify candidates based on temporal patterns of cell shape and appearance features.

Our method follows the spirit of the hybrid approach. It takes a phase contrast microscopy image sequence as input, and automatically outputs localized sub-regions in the sequence where mitosis occurred. As shown in Fig. 1, the algorithm consists of two steps. First, microscopy image preconditioning [7] and volumetric segmentation are utilized to locate spatiotemporal sub-regions in the input image sequence where mitosis potentially occurred. Then, a *hidden conditional random field* classifier [9] is applied to classify each candidate sequence as mitosis or not. These two steps jointly maximize recall and precision, achieving accurate detection. We will present the technical detail of each step in the subsequent sections, with emphasis on the second step.
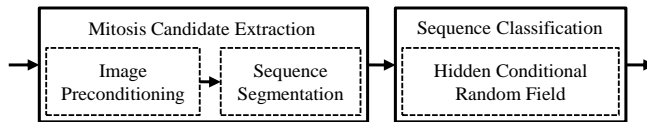

Fig. 1.   Mitosis Detection Workflow

[†] A-A Liu is currently with the Department of Electronic Information Engineering, Tianjin University, Tianjin, China.
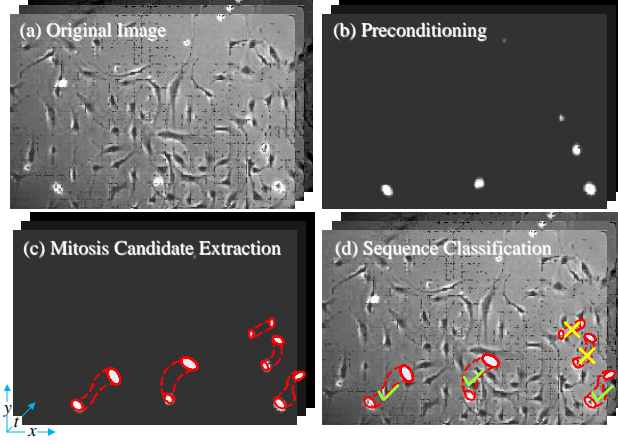[‡] K. Li is currently with Microsoft Corporation, Redmond, WA.

Fig. 2. Key Steps of the Proposed Method

2. MITOSIS CANDIDATE EXTRACTION

The mitosis candidate extraction step serves the purposes of eliminating "easily" negative regions in the image sequence where mitosis is unlikely to occur, and extracting temporally continuous sub-sequences with potential mitosis to facilitate subsequent sequence classification. The algorithm consists of two sub-steps. First, we apply the nonnegative mixed-norm algorithm proposed by Li *et al* [7] to precondition each input image. The algorithm leverages a phase contrast image formation model and transforms the input into an *ideal* image with zero background and nonzero foreground regions that correspond to potential mitotic cells (Fig. 2(b)). The image formation model is defined by an effective point spread function (or EPSF):

$$EPSF(x,y) \propto \delta(x,y) - (\alpha e^{-\frac{x^2+y^2}{\sigma_1^2}} - \beta e^{-\frac{x^2+y^2}{\sigma_2^2}}) \quad (1)$$

where $\delta(\cdot)$ is a Dirac delta function, and α, β are scaling factors. The EPSF approximates the imaging function of phase contrast optics, which accounts for the formation of halo effects around imaged cells. The ideal image is obtained by solving a linear inverse problem using an efficient multiplicative-update algorithm. We refer the interested readers to [7] for more details on the algorithm.

After preconditioning, 3D seeded region growing is applied to the transformed image sequences to extract spatiotemporal sub-regions that correspond to candidate mitosis sequences (Fig. 2(c)). The algorithm relies on two automatically-determined thresholds: a seeding threshold computed by Otsu's optimal thresholding algorithm is used to detect seeds; and a lower threshold determined by Rosin's unimodal thresholding algorithm [8] is used as the stopping criterion of region growing.

3. SEQUENCE CLASSIFICATION

The core of the sequence classification step is the *hidden conditional random field* (HCRF) classifier. We briefly review the basics of HCRF and two closely related models.

3.1. Hidden Conditional Random Fields

Generative dynamic Bayesian network models, in particular the hidden Markov model (HMM), are widely used for labeling sequential data. A limitation of such models is that observations are assumed to be independent given the values of hidden variables (*i.e.*, labels), which makes them unsuitable for incorporating long range dependencies between observations and their labels. This limitation leads to the introduction of discriminative models for sequence labeling, most notably the conditional random field (CRF) model [10]. A CRF model specifies the probabilities of possible label sequences given an observation sequence. The conditional dependency of each label on the observation sequence is specified through an arbitrary number of *feature functions*, and these feature functions can access the entire input sequence at any time during inference. These flexibilities enabled CRF to outperform HMM and become immensely popular for natural language part-of-speech tagging and biological sequence analysis.

A drawback of CRF is that it assumes the label sequence to be fully observable, and thus all frames in every training sequence must be fully labeled. This makes it inconvenient for sequence classification tasks in which each sequence is to be assigned a single label. To mitigate this drawback, Quattoni *et al* [9] proposed a hidden(-state) CRF model. HCRFs use intermediate hidden states to model the latent structure of the input domain, and infer a single label for an input sequence. This allows us to use training sequences not explicitly labeled frame-by-frame.

Mathematically, HCRFs deal with the problem of predicting a label y given an observation sequence $\mathbf{X} = \{x_1, x_2, ..., x_T\}$, where y is a member of a set $\mathbf{Y}$ of all possible labels. Each observation x_i is represented by a feature vector $\phi(x_i) \in \mathbf{R}^d$. For each sequence, we also assume a vector of hidden variables $\mathbf{h} = \{h_1, h_2, ..., h_T\}$, which are not observed in the training examples. A graphical representation of the HCRF model is shown in Fig. 3.
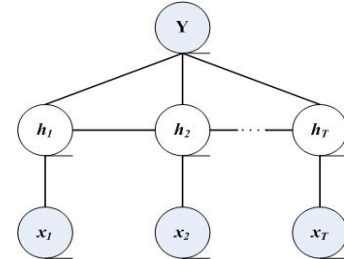

Fig. 3. Graphical Model of HCRF

Given the definitions of the label y, the sequence of observations $\mathbf{X}$, the hidden variables $\mathbf{h}$ and the model parameters θ, the HCRF model can be defined by:

$$p(y \mid \mathbf{X}, \theta) = \sum_h p(y, \mathbf{h} \mid \mathbf{X}, \theta) = \frac{\sum_h e^{\psi(y, \mathbf{h}, \mathbf{X}; \theta)}}{\sum_{y' \in \mathbf{Y}, h \in \mathbf{H}} e^{\psi(y', \mathbf{h}, \mathbf{X}; \theta)}} \quad (2)$$

where $\psi(y,\mathbf{h},\mathbf{X};\theta) \in \mathbf{R}$ is a potential function parameterized by θ as:

$$\psi(y,\mathbf{h},\mathbf{X};\theta) = \sum_{j=1}^{m} \sum_{l \in L_1} f_{1,l}(j,y,h_j,\mathbf{X})\theta_{1,l} + \\ \sum_{(j,k) \in E} \sum_{l \in L_2} f_{2,l}(j,k,y,h_j,h_k,\mathbf{X})\theta_{2,l} \qquad (3)$$

Here L_1 is the set of node features, L_2 is the set of edge features, $f_{1,l}$, $f_{2,l}$ are functions defining the features in the model, and $\theta_{1,l}$, $\theta_{2,l}$ are the components of θ, corresponding to node and edge parameters. The first type of feature function f_1 depends on a single hidden variable value in the model, while f_2 can depend on a pair of values.

The model parameters can be learned from training examples by optimizing the objective function [10]:

$$L(\theta) = \sum_{i=1}^{m} \log p(y_i \mid \mathbf{X}_i, \theta) - \frac{1}{2\sigma^2} \| \theta \|^2 \qquad (4)$$

where m is the total number of training sequences. The first term in the objective function is the data log-likelihood. The second term is the log of a Gaussian prior with variance σ^2. A gradient ascent algorithm can be used to search for the optimal model parameter $\theta^* = \arg\max_{\theta} L(\theta)$.

Given an unseen test sequence $\mathbf{X}$, the best corresponding label y^* can be computed by

$$y^* = \arg\max_y p(y \mid \mathbf{X}, \theta^*) \qquad (5)$$

In both HCRF and CRF models, we can incorporate long range dependencies controlled by a window size w. The parameter defines the amount of past and future observations to be used when predicting the state at time t ($w = 0$ indicates only the current observation is used).

3.2. Features for Classification

We extracted three different kinds of features from each frame of a candidate mitosis sequence:

- **Intensity Histogram** (IH, 5D), which describes the global distribution of pixel intensities;
- **Histogram of Oriented Gradients** (HoG, 144D), which captures the edge or gradient structure that is characteristic of local shapes [11]; and
- **Gist** (180D), which represents texture features that preserve local structural information [12].

4. EXPERIMENTAL RESULTS

The proposed method was validated in five challenging phase contrast image sequences of C3H10T1/2 mouse mesenchymal stem cell populations. The cells were observed under a Zeiss Axiovert 135TV inverted microscope, using a 5X, 0.15 N.A. objective lens with phase contrast optics. Images were acquired every 5 minutes for 120 hours using a 12-bit Qimaging Retiga EXi Fast 1394 CCD camera at 500ms exposure with a gain of 1.01. Each image consists of 1392×1040 pixels with a resolution of 19 μm/pixel. The relatively low resolution was chosen in order to image a large cell population in the limited field of view.

4.1. Performance of HCRF Classification

With preconditioning and volumetric region growing, we extracted candidate mitosis sequences in each input sequence. This step achieved 100% recall of detection with low precision. To improve precision, we used HCRF to refine the detection results. To train the HCRF model, we manually labeled all mitosis candidates in one sequence. The remaining four sequences were used for validation.

We trained HCRF models with IH, HoG, and Gist features and different window sizes. To choose the best configuration of features and window size, we plotted the ROC curve of each model and compared the area under curve (AUC) values. The results showed that the model trained with Gist features and $w = 2$ consistently outperformed the others with the best AUC value of 0.92. The ROC curves for the model using Gist features and a window size of 2 for four test sequences are shown in Fig. 4.
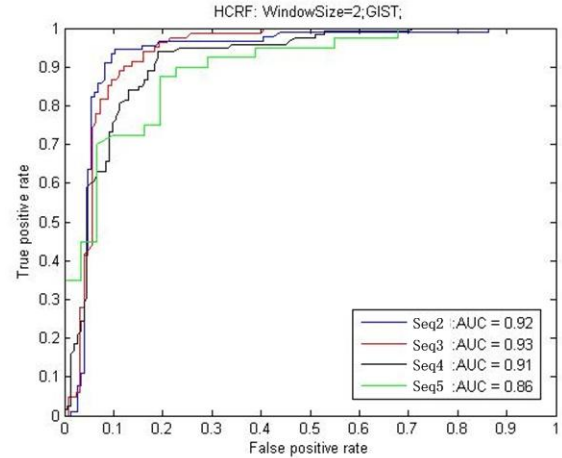

Fig. 4. ROC of HCRF with Gist Feature and $w = 2$

4.2. Comparison to CRF and SVM

To demonstrate the superiority of HCRF for sequence classification, we compared its performance to the CRF model trained with fully-labeled sequences. Moreover, to show the advantage of integrating temporal information, we compare its performance to a frame-by-frame classification approach using a support vector machine (SVM) classifier.

4.2.1. Conditional Random Field
To utilize CRF for sequence classification, it is first applied to label the full sequence. For training, we divided each mitosis sequence into four phases (Fig. 5), and assigned

labels 1 to 4 to each frame accordingly. Then, a candidate sequence is classified as mitosis if the number of frames assigned with labels 2 and 3 is greater than a threshold.

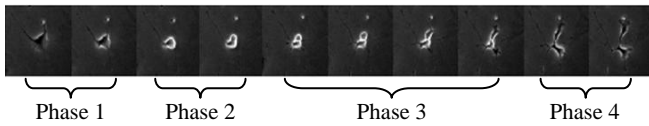Phase 1   Phase 2   Phase 3   Phase 4

Fig. 5.   Label for mitosis sequence

By varying the threshold, we obtained the ROC curves for CRF models trained with different features and window sizes. We found that CRF with Gist features and a window size of 2 achieve the best AUC of 0.78.

4.2.2. Support Vector Machine

The support vector machine (SVM) is a binary classifier that constructs a linear decision boundary (hyperplane) to optimally separate two classes [13]. We implemented a mitotic cell detector using SVM with a radial basis function (RBF) kernel. The detector was applied independently to each frame of a candidate sequence.

Corresponding to the training strategy for CRF, we labeled the frames that belong to phases 2 and 3 of a mitosis sequence as positive samples, and the others frames as negative samples. A candidate sequence is classified as mitosis if the number of frames assigned to be mitotic exceeds a certain threshold.

With cross-validation, we selected the best parameters for the SVM models trained with different features. By comparing the ROC curves for the trained models, we found that Gist outperformed the other features with the best AUC of 0.77, followed by HoG with 0.54, and IH with 0.47.

4.2.3. Overall Comparison

Finally, to compare the overall classification performances of HCRF, CRF and SVM with Gist features, we utilize the balanced *F* score as a complementary metric to AUC. The *F* score is defined as follows:

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

We separately computed the AUC and the best achievable *F* score for each sequence with each classifier. The results indicate that the HCRF classifier consistently outperformed both CRF and SVM, with a best-case performance of 95% precision and 85% recall ($F = 0.90$).

Table 1. Comparison of HCRF, CRF and SVM with Gist

Seq.	AUC			Maximum *F* Score		
	HCRF	CRF	SVM	HCRF	CRF	SVM
1	**0.92**	0.78	0.77	**0.90**	0.84	0.80
2	**0.93**	0.73	0.71	**0.87**	0.83	0.77
3	**0.91**	0.73	0.70	**0.86**	0.84	0.81
4	**0.86**	0.62	0.55	**0.87**	0.75	0.80

5. CONCLUSION

We proposed a fully-automated mitosis event detection method using hidden conditional random fields for cells imaged with phase contrast microscopy. The method consists of two stages, mitosis candidate extraction and sequence classification, which jointly maximize recall and precision. By experimentally comparing HCRF, CRF and SVM classifiers using intensity histogram, HoG and Gist features, we found that the HCRF model with Gist features achieved the best sequence classification performance. The method achieved 95% precision and 85% recall in very challenging phase contrast microscopy image sequences of C3H10T1/2 mesenchymal stem cell populations.

6. REFERENCES

[1] F. Yang, M.A. Mackey, F. Ianzini, G. Gallardo, and M. Sonka, "Cell segmentation, tracking, and mitosis detection using temporal context," *Proc. Med. Image Computing Comp. Assist. Interv.*, 8(1): 302–9, 2005.

[2] K. Li, E.D. Miller, M. Chen, T. Kanade, L.E. Weiss, P.G. Campbell, "Cell population tracking and lineage construction with spatiotemporal context," *Med. Image Anal.* 12(5): 546–66, 2008.

[3] L. Liang, X. Zhou, F. Li, S.T.C. Wong, J. Huckins, and R.W. King, "Mitosis cell identification with conditional random fields", *Proc. Life Sci. Sys. App. Workshop*, pp. 9–12, 2007.

[4] K. Li, E.D. Miller, M. Chen, T. Kanade, L.E. Weiss, and P.G. Campbell, "Computer vision tracking of stemness," *Proc. IEEE Int. Symp. Biomed. Imaging*, pp. 847–850, May 2008.

[5] B.A. Eccles and R.R. Klevecz, "Automatic digital image analysis for identification of mitotic cells in synchronous mammalian cell cultures," *Anal. Quant. Cytol. Histol.*, 8(2):138–47, Jun. 1986.

[6] G. Gallardo, F. Ianzini, M.A. Mackey, M. Sonka, and F. Yang, "Mitotic cell recognition with hidden Markov Models," *Proc. SPIE: Medical Imaging*, 5367: 661–8, 2004.

[7] K. Li and T. Kanade, "Nonnegative mixed-norm preconditioning for microscopy image segmentation," *Proc. Int. Conf. Information Processing in Med. Imaging,* Williamsburg, VA. July 2009.

[8] P. L. Rosin, "Unimodal thresholding," *Patt. Recog.*, 34(11): 2083–96, Nov. 2001

[9] A. Quattoni, S. Wang, L. Morency, M. Collins, and T. Darrell, "Hidden conditional random fields," *IEEE Trans. Pat. Anal. Mach. Intel.*, 29(10): 1848–53, Oct. 2007.

[10] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," *Proc. IEEE Int. Conf. Machine Learning,* pp. 282-89, 2001.

[11] N. Dalai and B. Triggs, "Histograms of oriented gradients for human detection," *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp. 886–93, 2005.

[12] Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," *Int. J. Computer Vision*, 42(3):145–75, 2001.

[13] V. Vapnik, Statistical Learning Theory, New York, Jon & Wiley, 1998.