SYMMETRY-BASED MITOSIS DETECTION IN TIME-LAPSE MICROSCOPY

Topaz Gilad*   *Mark-Anthony Bray*†   *Anne E. Carpenter*†   *Tammy Riklin Raviv**

*Electrical and Computer Engineering Department, Ben-Gurion University of the Negev
† Imaging Platform, The Broad Institute of Harvard and MIT

ABSTRACT

Providing a general framework for mitosis detection is challenging. The variability of the visual traits and temporal features which classify the event of cell division is huge due to the numerous cell types, perturbations, imaging techniques and protocols used in microscopy imaging analysis studies. The commonly used machine learning techniques are based on the extraction of comprehensive sets of discriminative features from labeled examples and therefore do not apply to general cases as they are restricted to trained datasets.

We present a robust mitotic event detection algorithm that accommodates the difficulty of the different cell appearances and dynamics. Addressing symmetrical cell divisions, we consider the anaphase stage, immediately after the DNA material divides, at which the two daughter cells are approximately identical. Having detected pairs of candidate daughter cells, based on their association to potential mother cells, we look for the respective symmetry axes. Mitotic event is detected based on the calculated measure of symmetry of each candidate pair of cells. Promising mitosis detection results for four different time-lapse microscopy datasets were obtained.

Index Terms— Mitosis detection, Symmetry, Time-lapse Microscopy, High-throughput images

1. INTRODUCTION

Mitosis is the process in which the previously duplicated genetic material in a cell undergoes nucleus division. The study of mitosis (or cell division) has a substantial impact on many fields in microbiology and biomedicine and cancer research, in particular. Time-lapse microscopy is an excellent platform to monitor cellular phenomena and events. As technology develops and acquisition techniques advance, biologists face enormous amounts of data. This leads to a growing demand for automatic tools for cell image analysis, as manual annotation becomes impractical.

1.1. Background and Related Work

Mitosis detection and cell tracking are intimately related. One can either use cell associations to detect cell divisions or use mitotic events as anchors for cell tracking. In [1] a mother cell is associated with one of the daughter cells, while the other daughter initiates a new track. Backward tracking is then performed to connect the two daughters with the mother cell. The main weakness of this strategy is the dependency of the mitosis detection quality on the tracking performances. Moreover, it requires an excessive computational complexity, when the purpose of the analysis is limited to the study of cell divisions. An alternative approach, if tracking is indeed needed, is to use the detected mitotic events which mark the initiation and the termination of cell paths to guide the tracking. As the appearance of dividing cells can be changed dramatically, current approaches use machine learning algorithms, such as support vector machine (SVM), to detect mitotic events based on the visual and temporal traits [2, 3, 4]. While these methods are successful when sufficient positive training examples are available along with a carefully selected set of distinguishing features, they cannot be applied to general cases due to the huge variability in cells' dynamics and appearance, caused by the differences in datasets, the chemical compounds used, the microscopy imaging technique and the imaging parameters.

1.2. Contribution

Consider the two pairs of consecutive frames, displayed in Figure 1, each showing mitosis of a different cell type acquired at a different rate. The proposed algorithm accommodates the difficulty of the different cell visual and temporal features without relying on the shape of the mother cell. Moreover, in contrast to most algorithms, it is not based on a full cell tracking, neither does it require a large amount of user-annotated data. Instead, as the majority of cells undergo symmetrical divisions, we detect mitosis by measuring the similarity between the two daughter cells right after the split. We note that this requires a sufficiently high frame rate to typically capture at least one frame of the cell when it is in anaphase or early telophase, the most symmetrical stages of mitosis. Using symmetry for mitosis detection has been suggested in [2], and applied to elongated cells, where the symmetry axis was determined based on the major axis of the mother cell. However, a symmetry axis is generally not as pronounced. Therefore, we suggest a novel algorithm to detect it based on our observation that the symmetry axis is orthogonal to the virtual straight line that connects the two daughter cells. We then quantify the level of symmetry based on gray-level as well as shape similarity, if a reliable cell segmentation is given. Encouraging detection rates were obtained for four different high-throughput microscopy datasets.

Frame t Frame $t+1$ Frame t Frame $t+1$

MFC-10A   N2DL-HeLa1
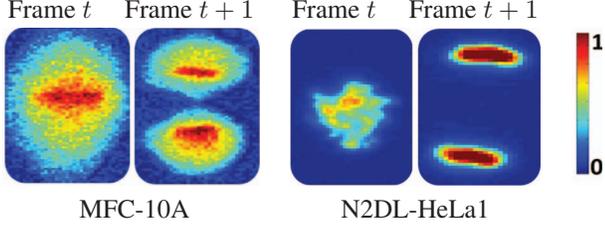
Fig. 1. Cell division from two different datasets: the mother cells are in frame t and the daughters in frame $t+1$. The cells of each pair are similar, although the datasets visually differ. Hereafter: 1. RGB colormap (instead of gray colormap) is used for highlighting. 2.MCF-10A dataset is courtesy of Albeck & Brugge. 3.N2DL-HeLa dataset source is [5, 6].

2. METHODS

The objective of the suggested method is a robust detection of mitotic events: Given a sequence of time-lapse microscopy images and their segmentation, a list of detected mitotic events is formed. The algorithm consists of two main stages: mitotic candidate extraction using mother-daughters relations, followed by examination of each candidate for mitosis by estimating a potential symmetry axis and calculating a similarity score between the candidate daughters. The score functions as a likelihood measure for being a mitosis.

2.1. Candidate selection

The goal of the first stage of the algorithm is to reduce the search space by creating a candidate list that will be further examined in the next stage for similarity between the candidate daughters. Therefore, that list should contain as much of the mitotic events, at the expense of additional false positives (FP). The construction of the candidate list is performed according to mother-daughters spatial proximity: Let $I^t : \Omega \to \mathbb{R}$ define an image frame acquired at time t where Ω is the $2D$ image domain. We define a candidate mother in frame t by $c_f^t, \quad f = 1, \ldots, F_t$ where F_t is the total number of detected cells in frame t. We search for candidate daughters in frame $t + 1$ in the region defined by $\omega_f^{t+1} \triangleq \{(x,y) \in \Omega | (x - x_f)^2 + (y - y_f)^2 \leq \mathcal{D}^2\}$ where (x_f, y_f) are the center of mass (COM) coordinates of c_f^t. The symbol $\mathcal{D}$ denotes a predefined search radius calculated from the sequence statistics, taking into account the average nearest neighbor within-frame distance and the average cell-drift between consecutive frames. The underlying assumption is that the COM of a mother cell (if not divided) would have been at the center of the line connecting the two daughter cells. We denote by $\{c_{f,d_i}^{t+1}, c_{f,d_j}^{t+1}\}$ the pair of daughter cells candidates which relates to c_f^t, if found and if none of them is closer to the COM of a different mother. If more than two cells are detected in ω_f^{t+1}, then the nearest pair is selected. For now on, the superscript $t + 1$ and the subscript f will be omitted for clarity.

2.2. Symmetry-based detection

Given a list of cell pairs, our objective is to detect the most likely daughter cells based on their similar appearance. Specifically, for each pair, we define a sub-image $I_{i,j} : \omega_{i,j} \to \mathbb{R}$, where $\omega_{i,j} \subseteq \Omega$, which contains the two daughter cell candidates and measure its degree of bilateral symmetry. We assume that if such symmetry exists, the symmetry axis is perpendicular to the straight line that connects the COM of the daughter cells candidates and intersects it at its center. We define the intersection point by (x_0, y_0) and the angle of intersection by θ_0. However, note that, as cell segmentation might be inaccurate the calculated COMs could be erroneous. We therefore re-estimate the pose and orientation of the symmetry axis, as will be described shortly, and use (x_0, y_0) and θ_0 for initialization.

Symmetry axis detection: We estimate the symmetry axis of a sub-image $I_{i,j}$ containing two daughter cell candidates in the spirit of [7]. Let $J_{i,j}$ be the symmetrical counterpart of $I_{i,j}$ obtained (w.l.o.g.) by an up-down flip. Let R define a planar rotation matrix and let $\tau = [\tau_x, \tau_y, 1]$ be a translation vector defined using homogenous coordinate system. We look for the Euclidean transformation $H(\tau_x, \tau_y, \theta)$:

$$H(\tau_x, \tau_y, \theta) = [R; \tau] = \begin{bmatrix} \cos\theta & \sin\theta & \tau_x \\ -\sin\theta & \cos\theta & \tau_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

that aligns $J_{i,j}$ to $I_{i,j}$. Formally, we solve the following optimization problem:

$$\{\hat{\tau}_x, \hat{\tau}_y, \hat{\theta}\} = \arg\max_{\tau_x, \tau_y, \theta} S_{sim}(I_{i,j}, H(\tau_x, \tau_y, \theta) \circ J_{i,j}), \quad (2)$$

where $J_{i,j}^H \triangleq H \circ J_{i,j}$ defines the transformation of $\omega_{i,j}$ (the image domain of $J_{i,j}$) by $H(\tau_x, \tau_y, \theta)$ and $S_{sim} \in [0, 1]$ is a similarity score to be defined next. Optimization was performed via Nelder-Mead algorithm[8]. As shown in [7], the orientation of the symmetry axis is $\theta/2$ with respect to the coordinate system of $\omega_{i,j}$, and its midpoint coordinates are $\tau_x/2, \tau_y/2$, given that the origin of $\omega_{i,j}$ is located at its center. The concept is illustrated in Fig. 2.

Similarity score: Let $D_{i,j}$ be the binary label map (segmentation) of $I_{i,j}$, obtained by setting (w.l.o.g.) the image pixels that belong to the cell image to one and the background pixels to zero. Let $U_{i,j}$ be the symmetrical counterpart of $D_{i,j}$ (the flipped image) aligned by $H(\tau_x, \tau_y, \theta)$. We define a similarity score $S_{sim} \in [0, 1]$ based on both the morphology and the gray level distribution of the two daughter cells candidates:

$$\begin{aligned} S_{sim}(I_{i,j}, J_{i,j}^H) &= \alpha_1 S_{corr}(I_{i,j}, J_{i,j}^H) + \alpha_2 S_{hist}(I_{i,j}, D_{i,j}) \\ &+ \alpha_3 S_{shape}(D_{i,j}, U_{i,j}), \end{aligned} \quad (3)$$

where, $S_{corr}, S_{hist}, S_{shape} \in [0, 1]$ are normalized similarity measures to be defined in the following and $\sum \alpha_i = 1$ are

Original image   Flipped image   Symmetry axis   Aligned image
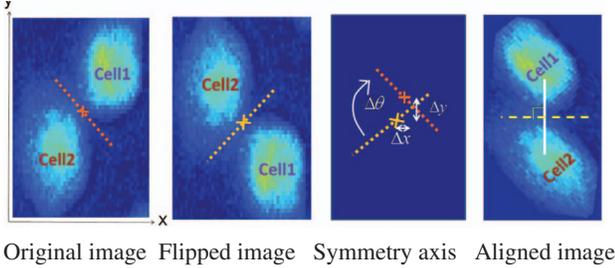
Fig. 2. Symmetry axis estimation. The symmetry axis is marked by a dashed red line. Its midpoint is marked with a cross. Its reflection is in orange. Parameters of the optimal H are in white. The original image is transformed by $\{\tau_x/2, \tau_y/2, \theta/2\}$ such that the symmetry axis is the horizontal axis that goes through the center of the sub image, and its midpoint coincides with the image center.

non-negative weights, selected such that the ratio between α_1 and α_2 is set to $1:3$ for all the datasets. We set α_3 to zero when the segmentation is not reliable.

Intensity correlation: The term S_{corr} could be obtained by a Pearson product-moment correlation between $I_{i,j}, J_{i,j}^H$. However, since, in dense environment, non-related cells can be captured in $I_{i,j}$ and break the symmetry, we use an approximated binary mask of the cells: $L_{i,j} = U_{i,j} \vee D_{i,j}$, where $\vee$ is the boolean operator OR. The score, S_{corr} is therefore the correlation result of $I_{i,j} \cdot L_{i,j}$ and $J_{i,j}^H \cdot L_{i,j}$ mapped to $[0,1]$. In the case where the label maps (segmentations) are not reliable, then a weighted correlation can be used instead:

$$S_{corr}(I_{i,j}, J_{i,j}^H) = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}, \qquad (4)$$

$$\sigma_{ij} = \frac{1}{N}\sum_{n=1}^{N} w_n(I_{i,j}^{(n)} - \overline{I_{i,j}})(J_{i,j}^{H,(n)} - \overline{J_{i,j}^H})$$

$$\overline{I_{i,j}} = \frac{1}{N}\sum_{n=1}^{N} w_n I_{i,j}^{(n)}, \quad \frac{1}{N}\sum_{n=1}^{N} w_n = 1.$$

In fact, the weights w_n are the values of two $2D$ Gaussian functions $G^i(o^i, \sigma_x^i, \sigma_y^i)$ and $G^j(o^j, \sigma_x^j, \sigma_y^j)$, where the origins o^i, o^j are the COMs of each of the daughter cells, and the variances σ_x, σ_y are calculated based on the major axes of the cells. Therefore, lower weights are assigned to pixels that are distant from the approximated cell's center as the confidence they belong to the cell is lower. Figure 3 demonstrates the main concept.

Histogram matching: Let h_i, h_j be intensity histograms of I_{d_i} and I_{d_j}, respectively, where, I_{d_i} and I_{d_j}, are the scalar product of $I_{i,j}$ with the label maps (segmentations) of daughter i and daughter j respectively. Let $l = 1, \ldots, \mathcal{L}$ define the bin indicies of the intensity histograms of I_{d_i} and I_{d_j}. As

(a)   (b)   (c)
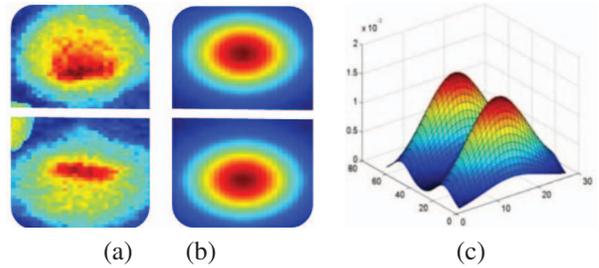
Fig. 3. Weights for correlation calculation: (a) Two daughter cells and additional cell that breaks the symmetry. (b) Colormaps representing w_n. (c) Two 2D Gaussians (mesh).

in [2], we define S_{hist} as follows:

$$S_{hist} = 1 - \frac{\sum_{l=1}^{\mathcal{L}} |h_i^l - h_j^l|}{\sum_{l=1}^{\mathcal{L}} \max(h_i^l, h_j^l)}, \qquad (5)$$

where h_i^l, h_j^l count the number of pixels in I_{d_i} and I_{d_j} (respectively) that their gray level values are within the range defined by bin l.

Shape similarity measures We define $S_{shape}(D_{i,j}, U_{i,j})$ by using the Dice score [9] which is a normalized measure of the overlap between $D_{i,j}$ and $U_{i,j}$.

3. EXPERIMENTS AND RESULTS

Experimental data: Four different datasets encompassing ten microscopy sequences were examined: 1. N2DL-HeLa1, see[5, 6] - HeLa cells stably expressing H2b-GFP, Olympus IX81 with plan 10x/0.4 objective lens, acquisition rate: 2.07 frames per hour (fph), resolution: 0.645×0.645 microns per pixel (mpp). 2. N2DH-SIM[1] [5, 6] - simulated nuclei moving on a flat surface, resolution: 0.125×0.125 mpp. 3. N2DH-SIM+[1] [5, 6]: Simulated nuclei of HL60 cells, rate 2 fph, resolution: 0.125×0.125 mpp. 4. A subset of MCF-10A[2] cells expressing NLS-mCerulean and RFP-Geminin, widefield epifluorescence, NikonTE2000, 20X objective, rate: 3fph and resolution: $0.2 - 0.3$ mpp. The obtained list of mitotic events was compared to annotations of mitotic events provided by the ISBI challenge [5, 6]. Cell segmentation of the MCF-10A dataset, which was not part of the challenge, was performed via CellProfiler [10]. For this dataset, we detected the mitotic events (for comparison with the automatic algorithm) by ourselves.

Results: A movie demonstrating our mitosis detection results for MCF-10A and N2DL-HeLa1 datasets is provided as a supplementary material. We considered only cell pairs that were fully within the frames. Table 1 shows the quantitative results, which include the true positives (TP), false positives

---

[1]Data source: V. Ulman and D. Svoboda, Centre for Biomedical Image Analysis, Masaryk University. Brno. Czech Republic (Created using Cytopacq).

[2]Data source: J. Brugge and J. Albeck

	#TP	#FP	#FN	Precision	Recall	F-Score
N2DH-SIM	104	5	12	95%	90%	92%
N2DL-HeLa1	87	15	7	85%	93%	89%
N2DH-SIM+	65	1	2	98%	97%	97%
MCF-10A	10	1	2	91%	83%	87%

Table 1. Mitosis detection results for four different data sets. The reader is referred to a movie demonstrating our mitosis detection results for MCF-10A and N2DL-HeLa1 datasets, which is provided as a supplementary material.
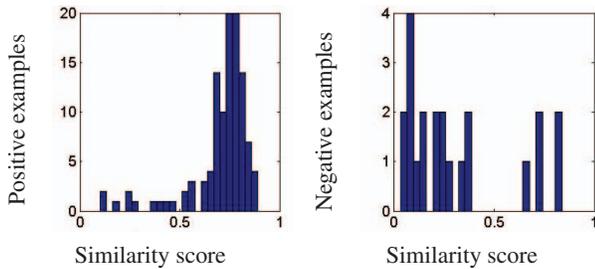
Fig. 4. Histograms of the similarity scores S_{sim} for the positive (left) and negative (right) examples in the candidates list of the N2DH-SIM dataset [5, 6]. Note that only very few false positives get high similarity score and therefore can be mistakenly detected as mitosis after the final stage.

(FP), false negatives (FN), precision, recall and F-scores, for each of the tested datasets, from the pre-selected list of candidates. Histograms of the similarity scores of the candidate list, obtained from the largest data set examined, are displayed in Figure 4. The similarity score S_{sim} allows good separation between the positive (mitotic events) and the negative cases. Nevertheless, there are a few examples of false detection (Figure 5). Addressing these outliers is discussed next.

4. SUMMARY AND FUTURE DIRECTIONS

We presented a novel method for mitosis detection which overcomes the huge variability of high-throughput microscopy imaging of cells. This is accomplished by addressing symmetrical cell divisions and using the detected symmetry axis as a guiding cue. High sensitivity and specificity measures were obtained for four different datasets of different sources.

Cells dynamics play an important role, distinguishing an actual cell division event from a random pair of cells passing by each other. Due to the forces of the microtubules in the mitotic spindle, pulling chromatin apart [11], the daughter cells appear as if they repel each other. This repulsion is captured within a frame or two after the splitting. Tracking the daughters, while measuring their relative distance and direction may provide additional useful information for classification.

It should be noted that many biological experiments involve cell perturbations which may cause asymmetric mitoses. If the process of candidate detection is highly accurate via cell tracking or by observing cell repulsion (as described above), the symmetry property can be used to highlight atypical cell

False Positive

False Negative

MFC-10A N2DL-HeLa N2DL-HeLa
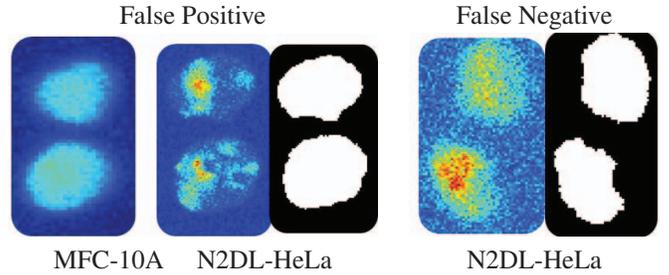
Fig. 5. Examples of false detection (left) and mis-detection in MFC-10A and N2DL-HeLa1 [5, 6] datasets.

divisions, which would be valuable to detect mutant phenotypes. Future work will aim to investigate these interesting directions, extending the scope of the proposed method.

5. REFERENCES

[1] K. Thirusittampalam, M.J. Hossain, O. Ghita, and P.F. Whelan, "A novel framework for cellular tracking and mitosis detection in dense phase contrast microscopy images," *IEEE J. Biomedical and Health Informatics*, vol. 17, no. 3, 2013.

[2] F. Li, X. Zhou, J. Ma, and S.T.C. Wong, "Multiple nuclei tracking using integer programming quantitative cancer cell cycle analysis," *IEEE TMI*, vol. 29, no. 1, pp. 96–105, 2010.

[3] S. Huh, *Toward an Automated System for the Analysis of Cell Behavior: Cellular Event Detection and Cell Tracking in Time-lapse Live Cell Microscopy*, Ph.D. thesis, Carnegie Mellon University, March 2013, PhD thesis.

[4] A.A. Liu, K. Li, and T. Kanade, "A semi-markov model for mitosis segmentation in time-lapse phase contrast microscopy image sequences of stem cell populations," *IEEE TMI*, vol. 31, no. 2, pp. 359–369, 2012.

[5] M. Maska et al., "A benchmark for comparison of cell tracking algorithms," *Bioinformatic*, vol. 30, no. 11, pp. 1609–1617, 2014.

[6] C. Ortiz-de Solorzan, "ISBI 2014 cell tracking challenge," 2014, [Online; accessed 19-July-2014].

[7] T. Riklin Raviv, N. Sochen, and N. Kiryati, "On symmetry, perspectivity, and level-set based segmentation," *IEEE TPAMI*, vol. 31, no. 8, pp. 1458–1471, 2009.

[8] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.

[9] L. Dice, "Measure of the amount of ecological association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[10] L. Kamentsky et al., "Improved structure, function, and compatibility for CellProfiler: modular high-throughput image analysis software," *Bioinformatics*, vol. 27, no. 8, 2011.

[11] D. Sadava, D.M. Hillis, H.C. Heller, and M.R Berenbaum, *Life: The Science of Biology*, 2012, 10th ed.