CrowdTrack: Interactive Tracking of Cells in Microscopy Image Sequences with Crowdsourcing Support‡

M. Sameki[*], **M. Gentil**[*], **D. Gurari**[×], **E. Saraee**[*], **E. Hasenberg**[*], **J. Y. Wong**[*], and **M. Betke**[*]

[*] Boston University, [×] University of Texas at Austin

Abstract. Outlining live cells and keeping track of their shape in microscopy videos are urgent tasks for biologists and medical researchers. As manual annotation by experts is time-consuming, automated solutions for specific types of cells or imaging modalities have been developed. We here propose CrowdTrack, a hybrid human-computer tracking method that can track various types of cells in fluorescence and phase-contrast microscopy videos by involving crowdsourcing whenever the performance of automated methods is unsatisfactory. We tested our proposed method on 1,523 frames from 12 different microscopy videos and obtained 14,351 cell outlines with only 32 5-worker rounds of crowdsourcing. CrowdTrack produced accurate cell outlines and correct cell lineage trees.

1 Introduction

How to effectively combine human and computer efforts to analyze image and video data is an interesting, relatively new research question. Prior work showed that crowdsourcing can scale up annotation of videos of every-day objects that are familiar to the crowd workers (Vondrick, Patterson, and Ramanan 2013). It has also been shown that crowdsourcing can be an effective tool for analyzing image content that is likely not familiar to the typical crowd worker: microscopy images of cells (Gurari et al. 2014; 2015). In this paper, we present the first tracking system that leverages the support of crowd workers who annotate videos of unfamiliar biomedical objects.

High-throughput microscopy technology enables researchers to produce large numbers of images of cells (Rittscher 2010) that developmental biologists use to analyze the life cycle and behavior of live cells. Manual annotation of these images, however, is costly and time consuming for experts, and so automated algorithms have been developed for cell segmentation (Chittajallu et al. 2015; Pan, Kanade, and Chen 2010; Song et al. 2013; Yin et al. 2015; Zhang et al. 2015) and tracking (Bise et al. 2009; Dzyubachyk et al. 2010; House et al. 2009; Meijering, Dzyubachyk, and Smal 2012; Wu et al. 2012). The algorithms can achieve acceptable performance for certain conditions but then fail to generalize in other conditions. In fact, the diverse appearance of different types of cells and recording difficulties pose a challenge for the development of a universal best method for cell tracking and segmentation (Chenouard et al. 2014).

One of the most challenging scenarios for tracking algorithms is cellular reproduction, also called mitosis. Mitosis is the process by which a cell duplicates its contents and then divides to yield two daughter cells with similar contents. Research has been conducted to address the task of tracking cells in this stage of their life cycle and its new born cells (Harder et al. 2009; Padfield et al. 2009; Held et al. 2010; Tsalik et al. 2012; Huh and Chen 2011; Huh et al. 2011). Padfield et al., 2009, for example, used level set segmentation to analyze cell cycle phases. Harder et al., 2009, and Huh et al., 2011, built a SVM classifier to predict cell cycle phases observed in fluorescence and phase-contrast microscopy sequences.

It would be challenging for a biomedical researcher without computer vision domain expertise to reproduce the automated cell tracking methods proposed in publications without accompanying code. We here propose a tracking method called CrowdTrack, accompanied by source code, that intelligently involves crowdsourcing support. CrowdTrack automatically involves crowd workers when mitosis or other tracking shortcomings are detected in order to ensure both correct lineage tracking or tracing (following cells and their descendants over time) and accurate tracking (maintaining cell outlines). Involving crowd workers enables us to discover false positive detections of mitosis in particular. We show that CrowdTrack is generalizable to diverse types of data by successfully applying it to different types of cells imaged with phase-contrast and fluorescence microscopy.

Our work makes a contribution to the methodology of crowdsourcing by showing that video annotation problems can be solved effectively and inexpensively by assigning inhomogeneous multi-round tasks to the crowd. In fact, CrowdTrack was able to extract a huge number of cell outlines (14,351) by employing only 32 5-worker rounds of crowdsourcing.

2 Method

2.1 Automated tracking

In order to segment cells for every frame, we employed popular level set methods. We chose a Chan-Vese Active

---
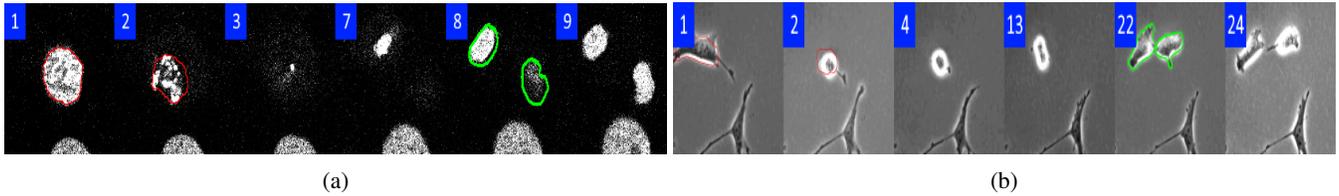
Figure 1: Interactive Mitosis Detection. Tracking results for cells undergoing mitosis in fluorescence (a) and phase-contrast (b) microscopy videos. The red boundaries surround the cell that is currently tracked until mitosis is detected in frame 2. The crowd workers were asked to follow the cell that has an overlaid red boundary and choose the frame where the newborn cells were visible. A new round of crowdsourcing produced the boundaries of the newborn cells (shown in green). Frame numbers (blue) were displayed to facilitate the workers' task. This example illustrates that mitosis detection is a challenging task for automated methods as mitotic cells display conflicting behavior for different microscopy modalities: in fluorescence images, the cell shrinks and decreases its luminosity (a), while cells in phase-contrast images (b) drastically reach a peak in pixel intensity when undergoing mitosis.
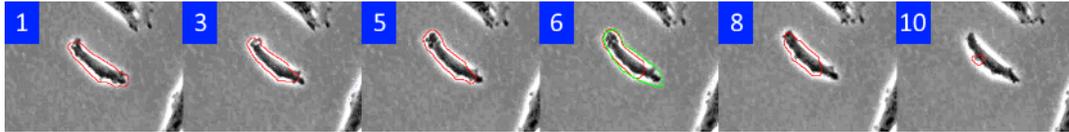
Figure 2: Interactive Correction of Cell Boundaries. Frames extracted from a video that was produced by CrowdTrack when it lost track of a cell. The red boundary surrounds the cell that was currently tracked until the mistake happened and the tracking was stopped. The crowd workers were asked to follow the cell that has an overlaid red boundary and choose the frame where the red outline is not accurately enclosing the tracked cell. A new round of crowdsourcing produced a corrected boundary for that cell, displayed as a green outline on the frame that was selected by workers. This example shows that non-expert crowd workers can recognize the time step where the automatic tracking fails and fix its shortcoming.

Contours approach (Chan and Vese 2001) to segment cells in fluorescence images (Fig. 1a) and the Caselles method (Caselles, Kimmel, and Sapiro 1997) for phase-contrast microscopy frames (Figs. 1b and 2).

Initial loose boundaries for cells in the first frame of each image sequence need to be provided to initialize the level set algorithms. This can be done by experts through the freely available LabelMe interface (Russell et al. 2008). Then the level set algorithm outputs a tight boundary for each cell in the first frame. In subsequent frames, first, dilation is performed on boundaries of cells in the previous frame to obtain initial contours for cells in the current frame, and then level sets are employed to iteratively tighten these contours around the cells.

When dealing with challenging datasets with highly variable luminosity and densely populated conditions, these algorithms are affected by mistakes that pose problems to an accurate tracking. In particular those cells that undergo mitosis need to be tracked carefully as they evolve their shape and luminosity very quickly and unexpectedly; furthermore mitotic cells exhibit diverse behavior for different microscope modalities and cell types, as shown in Figure 1. Therefore algorithms often fail to correctly keep track of the two newborn cells.

2.2 Classification for mitosis detection

For CrowdTrack to automatically detect when mitosis occurs, we built and trained a k-nearest-neighbor classifier with a parameter of $k = 3$ neighbors. CrowdTrack classifies each cell in each frame of our image sequence as either undergoing mitosis or not. In order to distinguish these two classes, CrowdTrack computes several features for each cell image region: circularity, area, and mean and standard deviation of pixel intensities (Harder et al. 2009; Huh and Chen 2011; Huh et al. 2011). It also uses the ratios of the features computed for the current frame over the mean value of the features observed during the previous frames. Ground-truth labels about mitosis occurrence are provided by an expert.

2.3 Tracking with crowdsourcing support

When automated tracking is started, features for each cell in each frame are extracted and fed to the classifier.

Crowd Task 1: Mitosis detection. When the classifier detects mitosis for two frames, CrowdTrack stops the tracking process and produces a video in GIF format showing only the cell during the mitosis event by cropping the image region in each frame around the place where the cell is initially situated (Figure 1). CrowdTrack then starts the first round of crowdsourcing, showing the video to five workers and asking them to select the first frame where both newborn cells are completely visible and remain so for the subsequent frames. Workers have the option to report that the problematic cell is not undergoing mitosis. If the majority of workers decides that mitosis is not happening, the classifier prediction is treated as a false positive situation; CrowdTrack then resumes tracking without any alterations.

Crowd Task 2: Outlining the newborn cells. If the majority of the crowd workers decides that mitosis is indeed

happening, the frame with the most votes is chosen to represent the start frame for the newborn cells. CrowdTrack plugs the selected frame into the LabelMe platform and summons a second set of five crowd workers to draw loose boundaries around each cell. The five drawings per cell are combined by "majority voting" into a single representation of the new cell as follows: a pixel is deemed to belong to the cell if at least 3 crowd workers included it in their annotation. The outlines of each newborn cell, created with this crowd support, are then used by CrowdTrack to initialize the level set segmentation algorithm and resume automated tracking.

Crowd Task 3: Detecting tracking mistakes Whenever CrowdTrack notices that a cell boundary is lost during tracking, it asks the crowd to fix the problem as follows. CrowdTrack creates a GIF video with the tracking output for that cell shown as a red overlaid boundary on these frames (Figure 2). Crowd workers are asked to follow the red outline for the cell and select the first frame where the red outline does not encircle the tracked cell properly. Majority vote on the frame number is again employed to increase reliability.

Crowd Task 4: Providing new boundaries The obtained cropped frame goes through one more round of crowdsourcing where all of the cells in it are annotated through the LabelMe platform. A new outline for each cell is produced by applying majority voting to the annotations provided by the workers as in crowd task 2. Hence, CrowdTrack can re-initialize the mistaken cell lineage by using the obtained boundary and rerun the level set methods.

Crowdsourcing design choice. Crowd tasks 1 and 2 could have been combined into a single crowdsourcing round. Similarly, crowd tasks 3 and 4 could have been done by the same workers. However, separating the tasks as proposed is advantageous for two reasons – efficiency and accuracy. It is more efficient to first determine a *single* representative frame for mitosis or a tracking mistake by majority vote. Otherwise outlines could be drawn on different frames and could not easily be combined for restarting tracking accurately.

3 Experiment and Results

3.1 Data Collection

We assembled a library of 20 videos of live cells. We used both phase-contrast and fluorescence microscopy image sequences. The phase-contrast images were collected in-house and show fibroblasts of a mouse strain recorded with a Zeiss Axiovert S100 microscope every 30 seconds, resulting in 3,897 frames. The fluorescence microscopy data were obtained from the 2013 Cell Tracking Challenge (http://www.codesolorzano.com/celltrackingchallenge), (Maška et al. 2014). Image sequences of four different kinds of cells include Chinese hamster ovary cells, mouse stem cells, rat mesenchymal stem cells, and simulated nuclei moving on a flat surface.

3.2 Experimental Methodology

We used the Amazon Mechanical Turk internet marketplace to perform our crowdsourcing experiment, hiring 5 workers

for each task. Out of 20 videos available, we selected the 12 we deemed most appropriate for automated tracking and manual annotation. The remaining 8 videos (which included the simulated nuclei) were only used for training purposes. We tested CrowdTrack in a leave-one-out round-robin fashion, where each of the 12 videos was tested on a system that was trained on 11 plus 8 videos. An expert provided ground truth labels about mitosis occurrence for 25,040 cell regions obtained from automatic tracking. The videos that CrowdTrack automatically created for crowd task 1 contained 17 to 27 frames, depending on the type of data set. For crowd task 2, 11 frames preceding the tracking problem were used. In our experiments, CrowdTrack processed 1,523 frames that each included 9.42 cells on average.

To evaluate our results for cell segmentation, we summoned an expert to provide ground-truth lineages for all the cells in 38 randomly selected frames of our dataset. We then computed the Jaccard index to measure the overlap ratio $\frac{A \cap B}{A \cup B}$ between the expert-drawn cell region A and the region B produced by the tracking algorithm.

To evaluate the ability of the tracking algorithm to detect the cells and follow them in time, that is, lineage tracing, we employed the metric "TRA" proposed by Maška et al., 2014. This metric measures the difficulty of changing the acyclic lineage graph generated by CrowdTrack and the ground truth graph by computing the number of basic operations that are needed to make these graphs identical.

3.3 Results

CrowdTrack produced 14,351 cell boundaries with the support of 51 unique workers and with Jaccard scores of 0.662 for phase-contrast images and 0.769 for fluorescence videos. These accuracy scores are high, considering that prior work reports 0.85 for expert agreement, e.g., Gurari et al., 2015.

CrowdTrack also produced correct lineage graphs, which we were able to verify for fluorescence image sequences, for which we had ground-truth lineage graphs: CrowdTrack produced an average TRA score of 0.8485 (where 1 is a perfect tracking result and 0 is completely incorrect).

The automated tracking process was improved by 32 rounds of crowdsourcing that were required by CrowdTrack whenever the automated tracking was unsatisfactory. In particular, 10 rounds were performed to pin down a mitotic event. Two of these situations were actually misclassified by our model as a result of sudden changes of the shape of a cell; however, crowd workers were able to recognize them as false positives and reported accordingly. Six cycles of crowdsourcing were requested as a result of a cell boundary being lost. Six rounds were necessary in order to provide new boundaries for newborn or mistracked cells as explained in Section 2.3.

4 Discussion and Conclusions

Our study shows that imperfect tracking methods can be improved with the help of crowd workers. It is particularly interesting that the crowd work was highly accurate although the workers were probably not familiar with the video content of moving cells. A similar observation was made by

Gurari et al., 2014; 2015 who found that crowd workers can segment cells accurately. Our study provides further evidence that crowdsourcing has the potential to help scale up the annotation process of scientific video data, as needed for research in developmental biology. We showed that crowd workers can reliably support the tasks of identifying true and false positive detections of mitosis, and fixing the tracking mistakes of automated algorithms.

Even though the task of automatic cell tracking has been explored by the research community for several years, state-of-the-art methods are still not sufficiently robust, i.e., their performance highly depends on the data employed (Chenouard et al. 2014). CrowdTrack is generic in the sense that its performance does not intrinsically depend on the employed data. We have shown this by applying it successfully to both fluorescence and phase-contrast microscopy image sequences and different types of cells.

By making our full source code and data freely available, we aim to stimulate open collaboration among researchers and inspire efforts towards this fascinating and challenging field of study. Our code is modular, so it is straightforward to plug in, for example, other tracking methods. In the future, we plan to collect additional datasets to train and test our model. We will incorporate additional automated methods to improve the segmentation of clusters of cells and mitosis detection. Finally, we will develop and test additional crowdsourcing methodologies to produce an accurate, efficient, and robust hybrid system.

References

Bise, R.; Li, K.; Eom, S.; and Kanade, T. 2009. Reliably tracking partially overlapping neural stem cells in DIC microscopy image sequences. In *MICCAI Workshop on Optical Tissue Image Analysis in Microscopy, Histopathology and Endoscopy (OPTMHisE)*.

Caselles, V.; Kimmel, R.; and Sapiro, G. 1997. Geodesic active contours. *IEEE Trans. Image Process.* 22(1):61–79.

Chan, T., and Vese, L. 2001. Active contours without edges. *IEEE Trans. Image Process.* 10(2):266–277.

Chenouard, N.; Smal, I.; De Chaumont, F.; Maška, M.; Sbalzarini, I. F.; Gong, Y.; Cardinale, J.; Carthel, C.; Coraluppi, S.; Winter, M.; et al. 2014. Objective comparison of particle tracking methods. *Nat. Methods* 11(3):281.

Chittajallu, D. R.; Florian, S.; Kohler, R. H.; Iwamoto, Y.; Orth, J. D.; Weissleder, R.; Danuser, G.; and Mitchison, T. J. 2015. In vivo cell-cycle profiling in xenograft tumors by quantitative intravital microscopy. *Nat. Methods* 12(6):577–585.

Dzyubachyk, O.; van Cappellen, W. A.; Essers, J.; Niessen, W. J.; and Meijering, E. 2010. Advanced level-set-based cell tracking in time-lapse fluorescence microscopy. *IEEE Transactions on Medical Image Processing* 29(3):852–867.

Gurari, D.; Theriault, D.; Sameki, M.; and Betke, M. 2014. How to use level set methods to accurately find boundaries of cells in biomedical images? Evaluation of six methods paired with automated and crowdsourced initial contours. In *MICCAI-IMIC*. 9 pp.

Gurari, D.; Theriault, D.; Sameki, M.; Isenberg, B.; Pham, T. A.; Purwada, A.; Solski, P.; Walker, M.; Zhang, C.; Wong, J. Y.; and Betke, M. 2015. How to collect segmentations for biomedical images? A benchmark evaluating the performance of experts, crowdsourced non-experts, and algorithms. In *WACV*. 8 pp.

Harder, N.; Mora-Bermdez, F.; Godinez, W. J.; Wnsche, A.; Eils, R.; and Ellenberg, J. 2009. Automatic analysis of dividing cells in live cell movies to detect mitotic delays and correlate phenotypes in time. *Genome Research* 19:2113–2124.

Held, M.; Schmitz, M. H. A.; Fischer, B.; Walter, T.; Neumann, B.; Olma, M. H.; Peter, M.; Ellenberg, J.; and Gerlich., D. W. 2010. Cellcognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nat. Methods* 7(9):747–754.

House, D.; Walker, M. L.; Wu, Z.; Wong, J. Y.; and Betke, M. 2009. Tracking of cell populations to understand their spatio-temporal behavior in response to physical stimuli. In *MMBIA*. 8 pp.

Huh, S., and Chen, M. 2011. Detection of mitosis within a stem cell population of high cell confluence in phase-contrast microscopy images. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 1033–1040.

Huh, S.; Ker, D. F. E.; Bise, R.; Chen, M.; and Kanade, T. 2011. Automated mitosis detection of stem cell populations in phase-contrast microscopy images. *IEEE Trans. Med. Imaging* 30(3):586–596.

Maška, M.; Ulman, V.; Svoboda, D.; Matula, P.; Matula, P.; Ederra, C.; Urbiola, A.; España, T.; Venkatesan, S.; Balak, D. M.; et al. 2014. A benchmark for comparison of cell tracking algorithms. *Bioinformatics* 30(11):1609–1617.

Meijering, E.; Dzyubachyk, O.; and Smal, I. 2012. Methods for cell and particle tracking. *Methods Enzymol* 504(9):183–200.

Padfield, D.; Rittscher, J.; Thomas, N.; and Roysam, B. 2009. Spatio-temporal cell cycle phase analysis using level sets and fast marching methods. *Med. Image Anal.* 13:143–155.

Pan, J.; Kanade, T.; and Chen, M. 2010. Heterogeneous conditional random field: Realizing joint detection and segmentation of cell regions in microscopic images. In *CVPR*, 2940–2947.

Rittscher, J. 2010. Characterization of biological processes through automated image analysis. *Annual Review of Biomedical Engineering* 12:315–344.

Russell, B. C.; Torralba, A.; Murphy, K. P.; and Freeman, W. T. 2008. LabelMe: A database and web-based tool for image annotation. *Int. J. Comput. Vision* 77(1–3):157–173.

Song, Y.; Cai, W.; Huang, H.; Wang, Y.; Feng, D. D.; and Chen, M. 2013. Region-based progressive localization of cell nuclei in microscopic images with data adaptive modeling. *BMC Bioinf.* 14:173–188.

Tsalik, E. L.; Jaggers, L. B.; Glickman, S. W.; Langley, R. J.; van Velkinburgh, J. C.; Park, L. P.; Fowler, V. G.; Cairns, C. B.; Kingsmore, S. F.; and Woods, C. W. 2012. Discriminative value of inflammatory biomarkers for suspected sepsis. *The Journal of Emergency Medicine* 43:97–106.

Vondrick, C.; Patterson, D.; and Ramanan, D. 2013. Efficiently scaling up crowdsourced video annotation. *Int. J. Comput. Vision* 101,(1):184–204.

Wu, Z.; Gurari, D.; Wong, J. Y.; and Betke, M. 2012. Hierarchical partial matching and segmentation of interacting cells. In *MICCAI*. 8 pp.

Yin, Z.; Su, H.; Li, E. K. M.; and Li, H. 2015. Cell-sensitive phase contrast microscopy imaging by multiple exposures. *Med. Image Anal.* 25:111–121.

Zhang, X.; Xing, F.; Su, H.; Yang, L.; and Zhang, S. 2015. High-throughput histopathological image analysis via robust cell segmentation and hashing. *Med. Image Anal.* 16:306–315.