Morphology-based Features for Adaptive Mitosis Detection of In Vitro Stem Cell Tracking Data

T. Becker¹,²,³; A. Madany¹,²

¹Institute for Neuro- and Bioinformatics, University of Lübeck, Lübeck, Germany;
²Graduate School for Computing in Medicine and Life Science, University of Lübeck, Lübeck, Germany;
³Fraunhofer Institution for Marine Biotechnology, Lübeck, Germany

1. Introduction

In the field of regenerative medicine and tissue engineering, the farming and production of adherently growing (adult) stem cell populations is a major task in the field of adult stem cell production used for drug discovery and in the field of regenerative medicine. To assess the quality of a cell population, a crucial event is the mitotic cell division: the precise knowledge of these events enables the reconstruction of lineages and accurate proliferation curves as well as a detailed analysis of cell cycles. To serve in an autonomous cell farming framework, such a detector requires to work reliably and unsupervised.

Methods: We introduce a mitosis detector that is using a maximum likelihood (ML) estimator based on morphological cell features (cell area, brightness, length, compactness). It adapts to the 3 phases of cell growth (lag, log and stationary phase). As a concurrent model, we compared ML with kernel SVMs using linear, quadratic and Gaussian kernel functions. All approaches are evaluated for their ability to distinguish between mitotic and non-mitotic events. The large, publicly available benchmark data CeTReS (reference data set A with >240,000 segmented cells, >2,000 mitotic events) is used for this evaluation.

Results: The adaptive (unsupervised) ML approach clearly outperforms previously published non-adaptive approaches and the linear SVM. Furthermore, it robustly reaches a performance comparable to quadratic and Gaussian SVM.

Conclusions: The proposed simple and label free adaptive variant might be the method of choice when it comes to autonomous cell farming. Hereby, it is essential to have reliable and unsupervised mitosis detection that covers all phases of cell growth.

Therefore, the tracking of in vitro cultured cells has shown to be a strong tool for the analysis of these populations. Only here, cell lineages can be extracted verifiably and, automated, with a high accuracy. For example, a manual cell tracking can be used to look for rare events in hemopoiesis in in vitro cell culture [1]. An automated version can even be used to determine the fate of a cell, i.e. the direction of differentiation [2]. The fundamental basis for all these experiments is the correct assignment of the cells relationship, which is determined by a cell division, a mitosis.

Though, the accurate detection of a mitotic event is not crucial only to the reconstruction of genealogical trees but also to the complete task of automated cell analysis. To solve this problem, there are two basic approaches: the first approach is based on cell tracking data. Using the tracked cells, mitotic events are extracted using the spatiotemporal information provided by the cell tracks. This approach has been used in several works to reconstruct mitotic trees in time-lapse image sequences [3, 4]. The second approach is tracking-free. Here, a mitosis is detected by a specific optical pattern: before a cell divides, it contracts and its three dimensional shape appears more spherical. Due to this behaviour, the cell shines very bright when observed using a phase contrast microscope. This pattern can be detected without knowledge of the cell positions in the image [5, 6] and therefore requires no cell segmentation. The main difference between these two approaches is the usage of an alternative characteristic to detect the mitotic event. Obviously, the tracking-free approaches are limited to phase contrast microscopy, while the first approach is able to deal with any input modality that allows for a robust cell detection (oblique illumination, dark field, differential interference contrast).

Beside technical varieties, cell tracking experiments typically differ in one major point: the biological context of the analyzed cell culture [7]. Therefore, most time-lapse experiments are performed using a low cell density; the end of the log phase or the stationary phase are not recorded.
simply as these phases are not very accessible for the experimentalist, and for the automated process aggravate the cell tracking problem. Thus, most tracking algorithms are evaluated for low confluency scenarios only. Beside the limited biological information of low cell density scenarios, only a small number of cells is observed during one experiment. Furthermore, the varying experimental setups make the evaluation and comparison of different cell tracking algorithms difficult and challenging task. For example, the evaluation of the data set used in [3] is limited to a small number of cells: the population growth from a population number of 4 cells to 43 cells (39 mitoses). Although the number of used cells is increased fourfold (160 cells) in the work of Debeir et al. [4], no detailed evaluation of the mitosis detection is given because of a missing reference data set.

However, Rapoport et al. [8] showed how a robust cell tracking algorithm allows for the extraction of biological relevant information: for all three growth phases are included during an experiment. In this scenario, cells are seeded at a low cell density (273 cells in the first image, i.e. 74 cells / mm²), show a strong proliferation (doubling rate between 12 and 20 h) and are tracked until the culture reaches a high cell density of 2343 cells in the last image (i.e. 637 cells / mm²). The growing number of cells (confluency) forces the cell morphology to dramatically change over time [8]. Especially for highly automated cell farming, which is a scenario that becomes more and more reachable [9], we need highly reliable cell tracking algorithms that provide robust performance for all three growth phases (lag, log and especially the stationary phase).

In the following, we use a freely available cell tracking reference data set to evaluate and improve a simple but effective morphology-based mitosis detection framework that constantly works well for all three growth phases. We systematically test four basic morphology features (cell area, cell length, compactness and brightness) for their applicability for mitosis detection. We systematically extend a global maximum-likelihood estimator [10] to an extended approach that adaptively toggles into three growth phase-dependent modes. We compare the performance of this estimator with that of kernel support vector machines [11] with linear (L-SVM), quadratic (Q-SVM) and Gaussian (G-SVM) kernel. As the SVMs additionally get a labeled set of training data for learning, we expect the SVMs to perform superior to the proposed approach. Surprisingly, the basic estimator clearly outperforms the L-SVM and reaches a performance close to that of Q-SVM resp. G-SVM.

2. Material and Methods

2.1 Data

For this work we used image data from the publicly available cell tracking reference set introduced by Rapoport et al. (CellRes, ftp://cetres.emb.fraunhofer.de/cetres/refs/for details see [8]). For each time point, the size of the observed area (i.e. the number of cells) is calculated. The confluency is defined as the percentage of the observed surface that is covered by cells. Again, once all cells in an image have been segmented, this parameter can easily be calculated by dividing the union of all cell bodies by the size of the observed area (i.e. the
image size). Note: often, the terms confluency and proliferation are used synonymously. Using our definitions, this would only be true if the cells cover a constant area.

These population parameters for the CeTReS data are shown in Fig. 2. It can be seen that the population was observed over all three phases of cell growth: the lag phase (first ten hours), the log phase (time period 10–40 h) and the stationary or plateau phase (time period 40 h–end). The mean cell area is shown in Fig. 2c. Roughly, the cell area stays constant at 650 μm² for the first 20 hours of the experiment. Afterwards, the cell area decreases linearly to 370 μm² in the last frame. In Fig. 1, three snapshots at the same position are illustrating the change of cell area and other morphological characteristics. Obviously, the cells are becoming smaller over time, intuitively correlated to the cell density in the culture.

2.3 Cell Segmentation and Tracking

To extract the spatiotemporal movement and the relationship for all observed cells, each single cell needs to be followed over time. This task consists of three different steps: first, we need a consistent input (pre-processing). Second, all cells need to be detected (cell segmentation) and third, the cells are followed over time.

2.3.1 Preprocessing

As a result of the good image quality, only one preprocessing step is needed to get a consistent input data for the segmentation algorithm: The background of the image data is estimated using a Gauss filtered image and is subtracted to correct for inhomogeneous illumination.

2.3.2 Cell Segmentation

To detect our desired objects, the cells, a global threshold segmentation is used. The global threshold is calculated using the method of Otsu [12]. The result of this operation is a binary image called the cell mask (Fig. 3d). Afterwards, morphological operations are used to delete small particles (cell debris and dirt) and to separate wrongly connected cells. Finally, all connected regions are detected as a single cell and are consecutively numbered in each frame (the i-th cell in frame t gets the unique label C_i,t). Our cell detection algorithm performs well with a false acceptance rate (FAR) and false rejection rate (FRR) of 0.05 and is comparable to other cell detection algorithms [13].

2.3.3 Cell Tracking

To find a successor cell, the intersection of two consecutive cell masks in frame t and t + 1 is calculated. If one representation of a cell (C_i,t) in frame t intersects (overlaps) with only one other cell in frame t + 1, both cells belong to the same spatiotemporal representation of one biological cell (Fig. 3e). Using this simple overlap tracking, each cell is followed over time starting at its first appearance. The cells are followed as long as they have one unique successor and are collected in so called “cell paths”. A cell path consisting of m cells is written as p = [C_n_t, C_{n-1}_t, ..., C_{m}{t}]. Here, n_t label the cell number and frame of the m-th cell of path p. If one segmented cell has more than one successor, the tracking is stopped and new cell paths are started at each successor cell.

At this point, all cells in the image series can be detected and followed over time. Our tracking algorithm creates a set of trustworthy cell paths, but leaves out the correct connection, which corresponds to the relationship of the cells, i.e. mother-daughter relations. Thus, it is not possible to extract complete lineages or to find simple relations like sibling cells in our data. To extract this information, it is essential to detect mitosis with a high accuracy.

2.4 Reference Data

To create a ground truth, the introduced cell segmentation and tracking algorithms were applied to the image data provided by the CeTReS data set [8]. Using the labeled reference data (CeTReS), the resulting cell paths were separated into two groups: first, all paths with a mitotic end were collected in one group; all remaining paths represent the non-mitotic group. This data was used as a ground truth.
truth for our mitosis detection, i.e. the task is to separate all paths that end in a mitosis from all paths that end in tracking irregularities.

2.5 Single Cell Morphological Features

Given the cell areas (binary masks) returned by our cell detection the following four morphological features were calculated for each cell in each time frame:

1. **Area.** The area $A$ of a cell is defined as the total number of pixels of the cell mask.
2. **Length.** To calculate the length $L$ of a cell, first an ellipse is fitted to the segmented cell. The length $L$ is then defined as the cell length measured along the ellipsoids main axis.
3. **Compactness.** The so-called compactness $C$ describes the circularity of a cell. It is defined as $C = \frac{4\pi A}{P^2} \geq 1$, where $P$ is the perimeter of the cell and $A$ is its area.
4. **Brightness.** To measure the brightness $B$ of a cell, the cell mask is used to calculate the mean image intensity in the corresponding patch from the original cell image.

For a single cell $i$ in frame $t$ all features are stored in the feature vector $f_i^t = [A_i^t, L_i^t, C_i^t, B_i^t]$, i.e. each cell is described using a four-dimensional vector. The matrix $P_k = [f_{n_1}^t, f_{n_2}^t, \ldots, f_{n_m}^t]$ is used to collect all cell features of path $k$, where $n_j, t_j$ describe the cell number and frame number of the $j$-th cell of path $k$, i.e. a cell path of length $m$ is described using a matrix with size $4 \times m$. The feature difference vector is denoted as $d_k^j = [d_{n_1}^j, d_{n_2}^j, \ldots, d_{n_m}^j]$, where $d_{n_j}^j = |f_{n_j}^t - f_{n_{j-1}}^t|$ stores the feature difference of cell path $k$.

2.6 Mitosis Detection

A mitosis (or cell division) is a major event during a cell’s life cycle and can be identified by its unique spatiotemporal representation when observed using time-lapse microscopy. The typical behavior of a cell during a mitosis is to stop moving and to contract: the cell gets smaller and rounder. One (technical) side effect of the decrease in size is typically an increase in brightness, which is caused by a different reflectance in the phase contrast. After all, the cell is not only contracting but also raising from the surface during mitosis. For all approaches proposed in the following, the last five entries of the cell path are used for mitosis classification (this is necessary because we are searching for the initial pattern of a cell division and the cells are tracked and followed during mitosis). Based on these patterns, we propose four different mitosis detection frameworks and compare their performance on the given benchmark data:

1. **Static Area and Compactness (SAC).** The first detection algorithm identifies mitotic data by comparing the values for area and compactness of a cell with a predefined threshold $\theta$. This detection algorithm is comparable to the algorithm proposed by Debeir et al. [4]. Please note that this algorithm does not work on the difference vector but rather on the absolute values of the features of the cell.
2. Single feature maximum likelihood (SFML). The SFML classifier uses feature difference vectors to test for a mitosis, similar to the detection introduced in [3]. The probability is estimated using the multidimensional normal distribution (\( \mu \) and \( \Sigma \) are the mean and covariance matrix of feature difference vectors of cell paths. To calculate \( \mu \), \( \Sigma \), the trustworthy paths returned by the automated cell tracking are used. To determine if a path is mitotic, the last five feature difference vectors are used to evaluate the probability \( \text{Prob}(d) \). If one of the probabilities falls below a predefined threshold \( \theta \), the path is labeled as being mitotic. In the single feature version (SFML), only one feature is selected to perform the mitosis detection, i.e. the multidimensional normal distribution is reduced to a one-dimensional normal distribution.

3. (3-Phase) Multi feature maximum likelihood (MFML and 3P-MFML). The MFML uses the same classification \( \text{Prob}(d) < \theta \) as introduced for the SFML. Both versions differ in the number of features that can be selected for the estimation: while SFML uses only one feature, the MFML classifier can be parameterized to any of the 15 possible combinations of the four features, in fact including the four SFML classifier. When all four features are selected for classification, this classifier is determined as MFML(static); this version corresponds to the mitosis detection used in [8]. In a second version, the set of static features is changed to a set of adapted features. For all three growth phases (lag, log and stationary phase) a different set of feature combinations can be chosen. This adapted classifier is called 3-Phase adapted MFML (3P-MFML).

4. Linear, Quadratic, and Gaussian Kernel Support Vector Machine (L-SVM, Q-SVM, and G-SVM). In general, a SVM aims to place a linear hyper plane between two given classes while trying to maximize the minimal margin, which is the distance from the closest point to that plane [14]. In an extension, the SVM is kernelized, i.e. the data is projected into a non-trivial kernel-space before placing the linear classifier, making the SVM much more powerful [15]. In the following, we use three SVM classifiers, one with a linear kernel function (L-SVM, i.e. a basic SVM), one using a quadratic kernel (Q-SVM) and one with a Gaussian kernel function (G-SVM) to classify paths as mitotic or non mitotic.

To create training data, all path ends labeled as mitotic (using the reference data) are used as positive samples (class mitotic), all other path ends are used as negative samples (class non mitotic). For each path end the entry with the biggest difference (measured as \( ||d||_2 \)) is extracted. Then, these feature differences are used to separate mitotic and non mitotic cells. For SVM training, a soft margin is used and all difference vectors \( d \) have been normalized (data is centered at its mean and scaled to have unit standard deviation).

2.7 Evaluation

To evaluate and compare the performance of the proposed classifiers, the receiver operating characteristic (ROC) and the area
under curve (AUC) were used. While the ROC curve allows to assess the global characteristic of a classifier by plotting the false acceptance rate (FAR) against the false rejection rate (FRR), we use the AUC to emphasize the strong impact of the changing cell morphology over time. Therefore, the AUC was calculated using a sliding window of \( t = 5 \) h (the size of the sliding window was chosen due to the small number of mitotic events in the lag phase (phase 1)). For the evaluation, a 10-fold cross-validation scheme was used; when the performance was tested using a sliding window or in each growth phase, only the path ends corresponding to the specific time window were used for cross-validation. The number of mitotic and non-mitotic events at each time step is given in Figure 5.

3. Results

Comparison of the ROC curves. The detection results for the SFML classifier using a single feature can be seen in Figure 6a. For comparison, the detection performance of the SAC classifier introduced by Debeir [4] is also shown in the same plot. The best SFML classifier is the one which is using the length of the cell as its solely feature. Surprisingly, the primitive SAC classifier, which classifies mitosis simply by the area and compactness of the segmented cell, still outperforms the SFML classifier for brightness, which is the worst performing SFML.

Although the compactness and length features describe a very similar characteristic of the cells, the SFML for length performs significantly better than the SFML that uses only compactness. The same holds true for the area: one would assume that the area of the cell is correlated to its length. But even though both SFML classifier for area and length show a very similar performance, the length still appears to be the strongest single feature that can be used for classification.

3.1 Comparison of AUC over Time

In Figure 6b, the SFML classifiers are compared over time. To estimate the performance we evaluated the AUC using a sliding window of 5 h for each single feature classifier. The data shows that again, the best performing feature is the cell length. It reaches an average AUC of about 0.83, while the cell size (area), which initially performs as good as the length, decreases significantly after the first 20 hours. This indicates that the change in area varies widely over time, especially compared for instance against changes in length. A classification with the SFML for compactness shows a contrary effect: during the first hours, this feature is not a good choice, but it starts getting better over time and finally even outperforms the cell area. But still, it does not reach the AUC that an SMFL classifier achieves for length.

3.2 3-Phase Adaption

To overcome the issue of changing detection performance over time, we propose to adapt the features for each of the growth phases individually. Therefore, we established the 3P-MFML classifier, which automatically switches to the optimal set of features in each phase. We used the time to determine the actual phase of cell growth. For automation, it is possible to identify the growth regime using the cell number (proliferation curve) or confluency.

---

**Fig. 6** Performance of different mitosis detectors. a) The ROC curve for the four SFML classifiers and for the (static) SAC approach is shown. b) The time-dependent strength is illustrated by the AUC applied on a sliding window of 5 h. Shown is the AUC for the SFML classifier (for all four features) and for its adapted version (3P-MFML). The three stages of cell growth are indicated by the vertical lines (phase 1 = lag phase, phase 2 = log phase, phase 3 = stationary phase).
To systematically find the best adaption in each phase, the ground truth data were split into three groups: the lag phase (0 to 10 h), the log (10–40 h) and the stationary phase (40–52 h). Due to the small number of features (4), we were able to check all possible combinations, resulting in the comparison of 45 AUC values. The feature sets having the highest average AUC score were then chosen for the specific phase. This resulted in the following list:
1. lag phase (0–10 h): [area],
2. log phase (10 h–40 h): [length compactness brightness],
3. stationary phase (40 h–52 h): [length compactness].

3.3 3P-MFML vs SFML

Naturally, the 3-phase adapted classifier (3P-MFML) performed better compared to the single feature SFML classifiers (see again Fig. 6a). But while area, brightness and compactness are clearly outperformed, the SFML classifier for length reaches a detection quality level surprisingly close to the adaptive one. The AUC curve (Fig. 6b) shows a similar behaviour for the 3P-MFML compared to the SFML for length. This is simply caused by the fact that the length parameter is chosen as one of the parameters the 3P-MFML classifier will adapt to in the log phase. The length and area features also show a rather similar SFML performance, thus it is not surprising that all three perform equally in this segment – after all, the adaptive algorithm picked area as its solely parameter for phase 1.

3.4 Support Vector Machine (SVM) classification (L-SVM, Q-SVM, and G-SVM)

The detection performance of different SVMs using linear (L-SVM), quadratic (Q-SVM) and Gaussian (G-SVM) kernel functions were assessed using a 10-fold cross-validation; for each kernel function three different SVMs were evaluated, one in each growth phase.

3.5 Comparison MFML vs. SVM

The performance of the MFML (MFML(static), 3P-MFML) and SVM (L-SVM, Q-SVM, G-SVM) classifiers are compared using the performance (AUC) calculated for each classifier in each growth phase separately (Fig. 7). The L-SVM performs worst in each phase (AUC [0.764 0.721 0.736]) while both non linear SVM perform similar (quadratic and Gaussian kernel function: AUC [0.856 0.876 0.884] and [0.862 0.882 0.889]), where the G-SVM sets the overall gold standard.

The 3P-MFML classifier performs superior to the MFML(static) in all three phases ([0.835 0.827 0.882] vs. [0.793 0.791 0.822]). The 3P-MFML reaches nearly the same performance as the Gaussian SVM in phase 3 (0.882 vs. 0.889), and is close to it in phase 1 (0.834 vs. 0.862) and phase 2 (0.827 vs. 0.882).

4. Discussion

Given an image series with all three growth phases observed, we were interested in the applicability of different morphological features (area, length, compactness and brightness) on the task of mitosis detection and the impact of confluency on this task. Therefore, the data was split into three groups: the main stages during cell culture: lag, log and stationary phase – which corresponds to major changes in cell behavior. The analysis of the single feature classification (SFML) shows that the change of length is the best feature to find mitoses during all three growth phases while the brightness is the worst of all tested features (Fig. 6). Both other features show a contrary behaviour: Figure 6 reveals that the (change of) compactness is a weaker feature to search for mitotic events than the size parameter in phase 1 (lag), while the compactness outperforms the area parameter at the end of phase 2 and in phase 3, which is correlated to the decreasing cell size over time (Fig. 2). To compensate for the changing detection performance, the introduced 3P-MFML classifier automatically adapts to the best feature in each growth phase and therefore outperforms each single feature (SFML, Fig. 6) and the classification using all parameters (MFML(static), Fig. 7) as used in
the non adaptive mitosis detection as proposed in previous work [8]. This result illustrates the impact of a confluency dependent adaption and the need for a detailed analysis of morphology-based mitosis detections over all three growth curves.

For comparison, SVMs with linear, quadratic and Gaussian kernel were tested. The Gaussian SVM gives the best detection accuracy in each growth phase and could be seen as the gold standard (Fig. 7), while the performance of the linear SVM indicates that the classes are most likely linearly inseparable.

A direct comparison of the 3P-MFML and the G-SVM classifier shows that the Gaussian SVM is superior (mean AUC 0.848 vs 0.878). Please note that two kinds of classifiers are compared: while the SVM uses a supervised training approach, the 3P-MFML classifier only depends on the automatically extracted path information and therefore is an unsupervised learning approach. The mean value μ and the variance σ of the difference features can be learned on unlabeled samples, only the optimal threshold θ is dependent on the FAR and FRR. Here, we were looking at the AUC, thus we tested for the general performance of the classifiers for all possible thresholds.

To use the MFML classifiers in practice, a suitable strategy is needed to learn a threshold. The confluency might serve as a reliable indicator on the number of mitotic events. To adapt the classifier, the appropriate parameters are simply turned on or off. Therefore, the proposed algorithm is most suitable on the way to fully autonomous cell farming systems.

As a gold standard, a supervised learning approach (i.e. labeled training data is necessary) was used: a support vector machine using a Gaussian kernel (G-SVM) was trained to classify mitotic vs. non mitotic cells. The G-SVM performs best with all kernels used, but remarkably the 3P-MFML classifier performs close to the G-SVM’s accuracy.

In a next step, these findings have to be evaluated systematically on different cell populations using varying tracking modalities to test for their universality.

5. Conclusion

We introduced a basic maximum-likelihood classifier (3P-MFML) for mitosis detection that automatically estimates the probability of a mitotic event based on the change of a given set of morphological features. The proposed algorithm adapts to the growth phase to overcome changes in the cultivated cell’s morphology, which are strongly correlated to the well-known three growth phases. All necessary parameters (μ, σ) are extracted during automated cell tracking, i.e. no labeled training data is needed. The current phase can be detected using simple measurements as the confluency or the cell number. To adapt the classifier, the appropriate parameters are simply turned on or off. Therefore, the proposed algorithm is most suitable on the way to fully autonomous cell farming systems.

A direct comparison of the 3P-MFML and the G-SVM classifier shows that the Gaussian SVM is superior (mean AUC 0.848 vs 0.878). Please note that two kinds of classifiers are compared: while the SVM uses a supervised training approach, the 3P-MFML classifier only depends on the automatically extracted path information and therefore is an unsupervised learning approach. The mean value μ and the variance σ of the difference features can be learned on unlabeled samples, only the optimal threshold θ is dependent on the FAR and FRR. Here, we were looking at the AUC, thus we tested for the general performance of the classifiers for all possible thresholds.

To use the MFML classifiers in practice, a suitable strategy is needed to learn a threshold. The confluency might serve as a reliable indicator on the number of mitotic events that can be expected on each frame. This estimate then could be used to approach a realistic threshold.

We hope that we can extend the CeTReS library soon with other pre-labeled data and that other groups might also contribute with their data. Then, a consequent next step would be to test on other time-lapse series how reliable and how reproducible the three phases of the 3P-MFML are working and how well the global parameters (μ, σ, θ) can be chosen completely unsupervised.

Acknowledgments

This work was supported by the Graduate School for Computing in Medicine and Life Sciences funded by Germany’s Excellence Initiative [DFG GSC 235/1].

References