COMP9517 Computer Vision Project: Cell tracking & Mitosis Detection

Chencheng Xie
Faculty of Computer Science and Engineering
University of New South Wales
NSW, Australia
z5237028@ad.unsw.edu.au

Gaoxiang Ren
Faculty of Computer Science and Engineering
University of New South Wales
NSW, Australia
z5216619@ad.unsw.edu.au

Jianqiu Chen
Faculty of Computer Science and Engineering
University of New South Wales
NSW, Australia
z5176348@ad.unsw.edu.au

Ruijun Zhou
Faculty of Computer Science and Engineering
University of New South Wales
NSW, Australia
z5223541@ad.unsw.edu.au

Zihao Li
Faculty of Computer Science and Engineering
University of New South Wales
NSW, Australia
z5176020@ad.unsw.edu.au

***Abstract*— This paper proposes a semi-automatic cell tracking algorithm containing cell detection, cell motion trajectory, mitosis detection, and motion analysis for various microscopic cell videos containing phase-contrast images (*PhC*), differential interference contrast (*DIC*), and fluorescent images (*Fluo*). The generalization of the model confronts issues like segmentation error, the mismatch between frames, and detection error for mitosis. To handle a variety of problems, we combined motion information and topological features to increase its robustness and accuracy among different datasets. Before segmentation, we applied a morphological closing filter to remove small noises and a top-hat filter to remove background shadings. Ultimate erosion followed by watershed is introduced for the segmentation of *Fluo* and *PhC*. As for *DIC*, because of the special characteristics of the cell images, deep learning model U-Net followed by a post-processing step test-time argument are used. In terms of tracking, we find the minimal cost regarding displacement, overlapping area, and deformation for the matching between consecutive frames. Eventually, the combination of intensity and topology change indicates the presence of cell division. The proposed method is experimented with several time-lapse sequences who are differ in illuminate condition, nonuniform cell motion, and demonstrates good performance on cell segmentation and moderate effect on cell tracking and mitosis detection.**

Keywords—Cell tracking, mitosis, watershed, time-lapse microscopy, phase-contrast image, fluorescence image, differential interference contrast image.

I. INTRODUCTION

Molecular biology develops gradually with more understanding of cells. In this case, cell tracking in microscopic videos is a critical step to allow biologists to get a further understanding of certain cells. The analysis of cell migration conveys a lot of information about cell action and related protein expression. Besides, the tracking of cells also indicates the information of mitosis. However, a normal approach for cell tracking is done by eyes, and it is quite tedious and may not be highly accurate [1]. With the development of technology, the automatic cell tracking method becomes more and more prevalent. Through automatic cell tracking, there are several parts including segmentation, tracking, mitosis detection, and analysis.

Segmentation is a crucial part of the whole process because the accuracy of the tracking step partially depends on the results of segmentation. Normally, there are several methods for cell segmentation, such as thresholding, mean shift, level-set, active contour, and watershed. Nevertheless, active contour requires the initialization of the objects [2], which is time-consuming and may introduce errors in the beginning. Mean shift and level-set are both commonly used in related cell tracking researches, but both methods need quite a long time to converge. Regarding the watershed segmentation algorithm, the main drawback of it is over-segmentation because of noises and multiple local maxima [3]. Besides, deep learning could also provide great performance at segmentation.

The mitosis incurs a manifest change in the cell life cycle, and thus it is important to discover the whole stage of mitosis while tracking. Cell migration not only comes with motions but also includes topological changes and mitosis. Considering these factors, traditional tracking approaches such as Kalman filter and particle filter may not be suitable for cell tracking because they work well for tracking with uniform motion [4]. Thus, the combination of several methods could be taken into account when it comes to different datasets with different migration motions and shapes.

In this paper, we proposed to present a partially automatic cell tracking method to do segmentation, cell tracking, and cell division detection. Because of the artifacts, low signal-to-noise ratio, different compactness, and various cell shapes, automatic tracking cells from different datasets may confront lots of challenges [1]. Thus, we decided to use watershed combined with some morphological preprocessing methods to tackle the segmentation problem for certain datasets and utilize a deep learning model to solve different problems for another one. Followed by that, a cost matrix combing several factors related to topological features and motion features was used to track the cells and build a trajectory between consecutive frames. Cell mitosis is also tackled by analyzing the information of cells such as circularity and intensity. Additionally, information regarding cell motions such as migration velocities, migration distances, and confinement ratios was obtained.

By proposing the method described in this paper, we can track most of the cells in consecutive frames as well as to

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

detect the whole phase of mitosis, and the cell motion analysis is also available.

The rest contents of this paper are structured as follow. Section II reviews the literatures in the field of computer vision and cell-tracking techniques. The designed method is explained in Section III. Section IV records the environment setups to conduct the method mentioned in this paper. Section V demonstrates the outcome of the proposed method on several sequences of microscopy and discusses the advantages and disadvantages of the proposed method. Lastly, Section VI concludes the accomplishments of this research and some suggested study directions.

II. Literature Review

To fully study the developments of automated tracking and movement in time-lapse microscopy image sequences, Meijering et al. [5] addressed the idiosyncrasy of the latest trends and successes in the tracking methods for molecules, cells, and organisms. For the main interests on the cell level, this paper compared the pros and cons of different approaches from simpler methods to more sophisticated methods. In the aspect of Segmentation on the individual cells: 1) thresholding is commonly used due to its simplicity, but it is also prone to error which brings the needs of more sophisticated techniques. 2) Fitting a predefined cell intensity profile is another approach, it performs well for cells with consistent shape but becomes problematic with cells whose morphology changes dramatically over time. 3) The watershed transform is a more widely used approach except that it is sensitive to noise and might lead to an over-segmented result. 4) The parametric contours and level-set functions were also introduced as the deformable model. The level-set function can easily capture the topological variance. This technique, on the other hand, tends to under-segment the images, and therefore, postprocessing is imperative to corporate the deformable model. In the field of Connecting cells over time, this paper stated that the straightforward approach is to associate the nearest cell in the next frame for each segmented cell. But for sequences with rapid cell-motion, the interpretation of "nearest" may extend to the feature space of eccentricity, area, perimeter, intensity, boundary curvature, etc. This paper concluded that there are no "jacks of all trades" to tracking all problems in cell and developmental biology. Dedicated solutions are needed for every specific problem.

Sacan et al. [6] developed a cell tracking and segmentation algorithm by implementing based on snakes for cell boundary detection and considering displacement and object rotation for tracking. To obtain the objects, snakes capture the contour of objects by computing minimal cost which takes both internal, external, and matching energy into account. It is indicated that by calculating the matching energy from the previous term, it reduces the possibility of under-segmentation and match the objects to its relative object in the previous frame. This term also associates the objects across frames which is useful in the tracking stage as well. They also used the ensemble method to combine various tracking algorithms such as template matching and optical flow in order to tackle the problem of deformations and object displacement. Besides, this application also takes advantage of image processing methods such as smoothing, and background subtraction to increase accuracy.

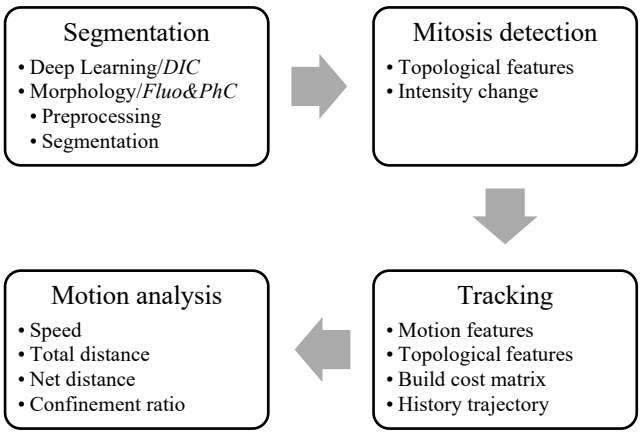


Fig. 1. Flow chart of the proposed model.

A cell tracking framework was proposed for phase-contrast microscopic videos [7]. The whole method was divided into two parts, the first part obtained a rough segmentation by analyzing object motions information such as velocity based on the theory of optical flow and the grey level changes. Then level set function using Chan-Vese was applied followed by active contour to refine the objects. Since level set function may alter the topology of objects while tracking so that causes over-segmentation or under-segmentation, so it was necessary to preserve the topology of each object. In order to preserve the topology, the initial contour of the objects was required to constrain the segmentation, and after the rough segmentation, the approximation object position could be obtained and refined using the active contour. [7] has shown a semi-automatic cell tracking method with strong robustness, however, this method works better for videos with reasonable duration, and for a longer duration, it may result in larger variance due to the deviation of the objects from its initial contour. To handle this problem, users may need to reinitialize the objects after an interval.

III. Methods

The whole method proposed in this paper includes 4 parts: cell segmentation, cell tracking, mitosis detection and motion analysis, the whole flow is shown in Fig. 1.

Before all main progress, since the given sequences may have certain sorts of noises such as background shading, preprocessing to some extent determines the performance of following segmentation and tracking steps. When it comes to the tracking stage, both topological features and motion features were taken into account [8], this combination contains moving distances, overlapping areas, and eccentricity of deformation.

A. Preprocessing

The types of cell images that fed into the automated processing pipeline were Fluorescence microscopy and Phase-contrast microscopy. The Fluorescence microscopy image brought challenges due to its low-contrast and possible dark noise which is rare in phase-contrast microscopy. The black top-hat filter, which involves morphological closing operation, was utilized to solve this issue. The dilation part of this closing operation dilates the local minima which have a smaller size than the structuring element. Since the dilation operation

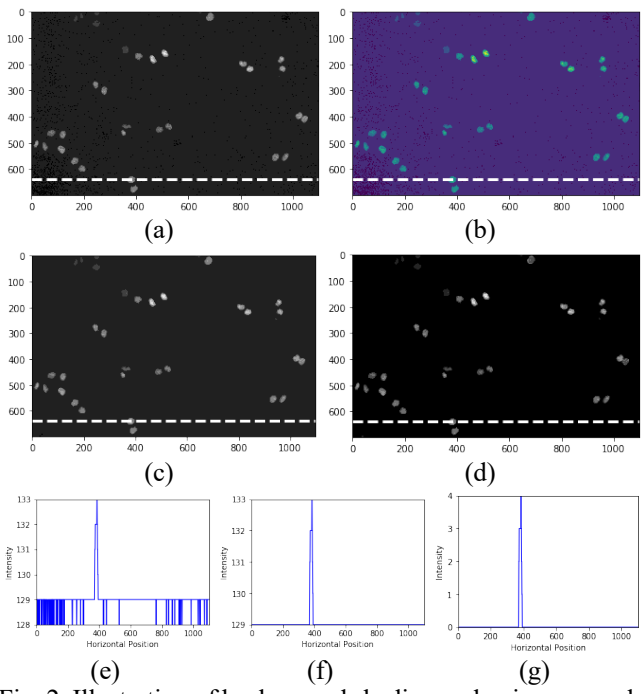


Fig. 2. Illustration of background shading and noise removal. (a) Input image of fluorescence microscopy. (b) Input image in viridis colormap. (c) Dark noise removed image of (a) using black top-hat filter. (d) Background shade removed image of (c) using white top-hat filter. (e)-(g) Horizontal profile of intensity for (a), (c)-(d), respectively, at row 640.

would enhance the brightness of the entire image, the erosion operation is introduced to adjust the brightness level. The filtered image would have the dark noise removed. After applying the black top-hat filter, the white top-hat operation can be conducted to remove the bright shades. Similar to black top-hat, white top-hat erodes out small local maxima and dilates the image to adjust the brightness level. The resulting image of opening operation is subtracted from the black top-hat filtered result. The purpose behind this subtraction is to transform the background to uniform distribution and enhance the dynamic range between the foreground and the background. We do not want to apply a black top-hat filter on the phase-contrast microscopy due to the preservation of the intensity difference between the cell and its surrounding halo. The mentioned procedure can be represented in one of the two following formulae:

for fluorescence microscopy:

$$\hat{I} = I - white\ top_{hat(black\ top_{hat(I,S)},S)} = I - \big((I \bullet S) \circ S\big) = I - max_S\{min_S[min_S(max_S(I))]\}$$

or, for phase-contrast microscopy:

$$\hat{I} = I - white\ top_{hat(I,S)} = I - (I \circ S) = I - max_S\big(min_S(I)\big)$$

where I is the input image, $\hat{I}$ is the processed image, and S is a square flat structuring element with edge length l. The edge length parameter of the white top-hat filter l_w is the size of cells by visual inspection, and of the black top-hat filter l_b is the size of noise. The Gaussian filter is then applied to further reduce the bright noise.

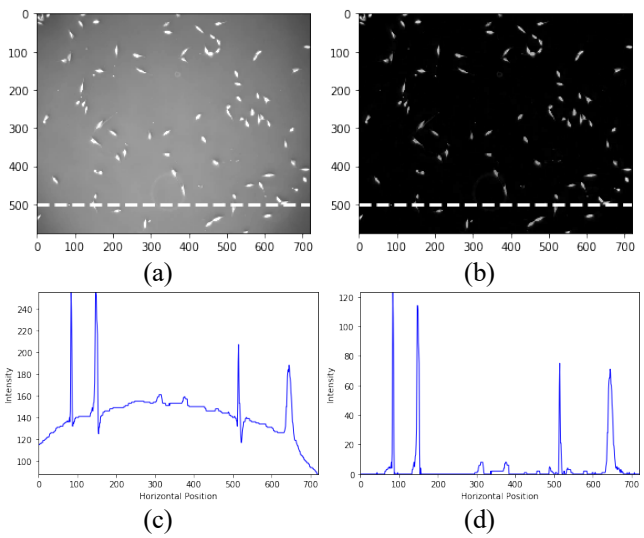


Fig. 3. Demonstration of background shading and noise removal. (a) Input image of phase-contrast microscopy. (b) Bright shade removed image of (a) using white top-hat filter. (c)-(d) Horizontal profile of intensity for (a)-(b), respectively, at row 500.

Fig. 2 and Fig. 3 demonstrates the effects of utilizing the black and white top-hat filters. Fig. 2(a) and (b) show the fluorescence image in grayscale and Viridis colormap, respectively. It might be worth mentioning that on the bottom-left corner of Fig. 2(b), a cluster of noises exist darker than the background. This was treated with black top-hat filter results in Fig. 2(c). By comparing the Fig. 2(e) and (f), it can be easily observed that all those regions with lower intensity in contrast with the background are removed. A white top-hat filter was then applied to adjust the background to zero in Fig. 2(d) and (g). It's even more clear in Fig. 3 how white top-hat filter suppressed the bright shade. The uneven illumination in Fig. 3(a) was also represented by the curvature of intensity in Fig. 3(c). This bright shading was fixed by applying the white top-hat filter, and the background intensity is uniformed and flattened in Fig. 3(b) and (d), respectively.

B. Segmentation

After the preprocessing, we obtained a sequence of microscopy images with even background and higher-intensity cells. But to actually separate the foreground and background, we need to apply binary thresholding to transform foreground pixels into 1s and background pixels into 0s. An automatic intensity thresholding algorithm [8] was adopted for phase-contrast microscopy to obtain the binary image. This algorithm exhaustively searches for thresholding value t that maximizes the variance between classes which are foreground and background in this case. This can be formulated as:

$$\max {\sigma_B}^2 = p_0 p_1 (\mu_0 - \mu_1)^2$$

where σ_B is the inter-class variance to be maximized, p_0 and p_1 are the fractions of pixels of foreground and background, μ_0 and μ_1 are the mean pixels intensities of foreground and background, respectively. For the fluorescence image, manual inspection on the intensity histogram and decision on the threshold value is imperative. This is due to the significant difference in brightness between the fluorescent cells and background. And to once again remove local peaks with

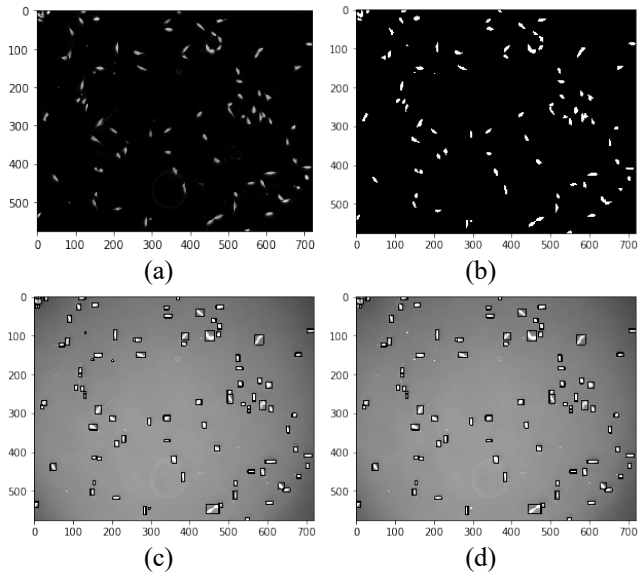


Fig. 4. Illustration of binary transformation and cell labelling. (a) Filtered image of phase-contrast microscopy. (b) Binary transformed image of (a) using the Otsu's thresholding value [6]. (c) Segmented cells image using the Watershed method. (d) Final image with small labels removed of (c).

intensity greater than thresholding value, we removed all connected components with the number of pixels smaller than a thresholding size τ. This thresholding size parameter τ should be determined by the smallest possible cell size. The binary image with small components removed was treated as the mask of the Watershed method [9]. And the seed of the Watershed method was obtained by the ultimate erosion method. The ultimate erosion method will iteratively erode each object right before it was removed, so each object will be left with the center points. But one issue brought with this ultimate erosion method is that it tends to over-segments the image and therefore increases error. The strategy we employed to deal with this over-segmentation problem was to remove the labels with sizes smaller than a threshold value φ. As we can possibly imagine, a label with a larger size indicates the further distance between the neighboring labels, and less likely to be an over-segmentation.

Fig. 4 demonstrates how filtered images are segmented. Fig. 4(a) shows a filtered phase-contrast image as described in the preprocessing section. Fig. 4(b) are binary transformed from (a) with Otsu's thresholding method. The small objects are also removed from Fig. 4(b). Fig. 4(c) are segmented results using the Watershed method with seeds from the ultimate erosion method. Small noisy labels can be observed from Fig. 4(c) and these noisy labels are removed in Fig. 4(d).

The differential interference contrast (*DIC*) dataset are trickier not just in term of low contrast between the interior and exterior, lots of details on the organelles are reserved, many of them behave cohesively and overlap with each other. The workflow of Gaussian filter, Total variation denoising [10], Contrast limited adaptive histogram equalization (CLAHE) [11], and *Meijering* neuriteness filter [12], *Otsu's* threshold [8], small object removal, and binary file holes was attempted. The processed image of first 6 frames of *DIC* was shown in Fig. 5. And it demonstrates that this method did not generalized well before further adjustments are studied.

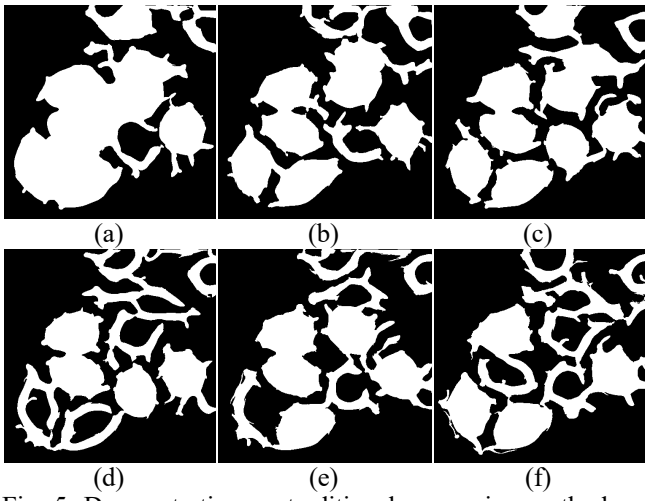


Fig. 5. Demonstration on traditional processing method on *DIC*. (a)-(f) The processed resulting image of frame 0~5.

Therefore, we abandon the traditional processing method on the *DIC* dataset, and instead we adopted the Deep Learning method.

The Deep Learning architecture we chose the U-Net architecture which is productive when the network only distinguishes between the foreground and background. U-Net is an architecture for processing biological images, still efficient in small training datasets, which is composed of two main parts. (1) a down-sampling convolutional network or encoder, through which we obtain a feature representation of the input image, and (2) an up-sampling convolutional network or decoder, which generates the segmentation from the feature representation of the input image [13].

Therefore, we selected the efficient model of MU-Lux-CZ which is effective in datasets [14]. The MU-Lux-CZ model generates two outputs, one called Predicted Markers Ω_1 and one called Predicted Cell Boundaries Ω_2. The prediction error for each marker is calculated based on the weighted cross-entropy as follow:

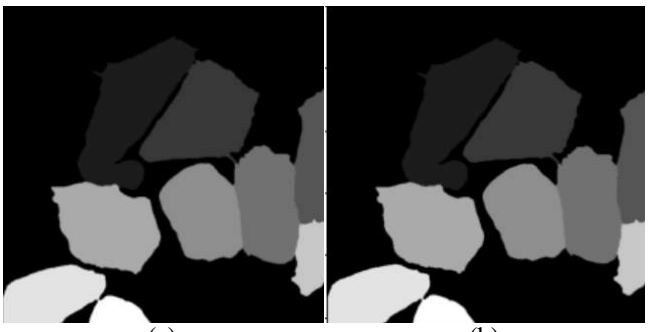
$$E(\Omega) = -\sum_{i\in\Omega} w_i log\left(p_{i,L(i)}\right)$$

where the weight parameters w_i of pixel p_i are trained by model mentioned above.

The resulting loss function is calculated by normalizing both layer average error of the sum of weights and can be formulated as follow:

$$L = \frac{1}{2} \times \frac{E(\Omega_1) + E(\Omega_2)}{\sum_{i\in\Omega_1} w_i + \sum_{j\in\Omega_2} w_j}$$

In the display area of the original image may be missing some important features. For improving the performance of results of segmentation, we perform post-processing with a Test-time Argument (TTA) after MU-Lux-CZ [3]. TTA includes four procedures: augmentation, prediction, dis-augmentation, and merging. We first apply the same augmentations on the different test images by generating multiple different versions of the test images to rotate as R_{angle}, transpose T, horizontal flip, vertical flip, etc. Then, we make a prediction on the original image and the image after



(a) (b)

Fig. 6. *DIC-C2DH-Hela* Segmentation in DL model. (a) is results based on model using MU-Lux-CZ. (b) is an image of post-processed TTA based on Fig. 8(a)

TTA, and then revert the prediction by return the image to its original orientation. Finally, the different versions of the images are merged to get an average output. This is used as the final output of the images.

TTA affects the borders of the object, but in the case of uncertainty, it can help to fit the whole object. In general, TTA improves segmentation accuracy in most cases. Although TTA requires multiple predictions for the same object at a higher cost, the analysis of uncertain regions in segmentation with our lower prediction base cost, and the post-prediction calculation with multiple objects makes the image more accurate and more robust. The Fig. 6 shows the before and after of TTA on U-Net resulting image.

C. Cell division detection

In the real cycle of living cells, cell division (mitosis) typically involves several steps including prophase, metaphase, anaphase, and telophase, resulting in the formation of two new nuclei each having the same number of chromosomes as the parent nucleus.

Before two new cells produced by mitosis are detected, we are supposed to determine which feature shows the start of cell division in order to represent a whole process of mitosis. In the tracking module, each bounding box containing information on the width and height of the cell has been drawn. It is observed that before a cell tends to divide into two new cells, the brightness within a cell is stronger than other normal cells and that cell tends to be circular. In this case, the mean intensity and circularity within each bounding box that we have done in the tracking module of the cell could be two factors to determine whether the process of cell division starts.

The mean intensity of each cell is calculated as follows:

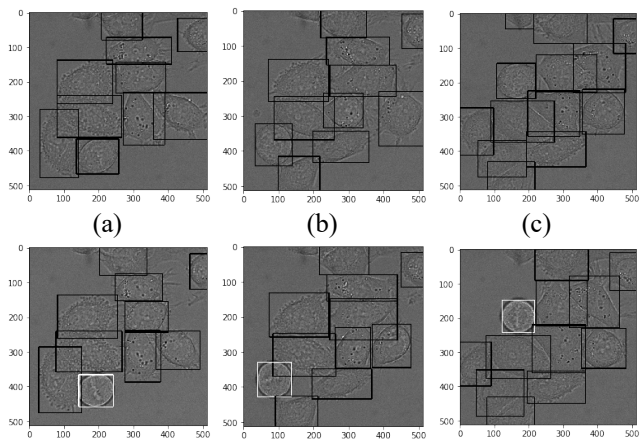
$$E_{intensity} = \frac{1}{m}\sum_{i=1}^{m} P_i$$

where m refers to the total number of pixels within the bounding box and P_i refers to i[th] pixel intensity value of pixels within bounding box.

The circularity of each cell is calculated as follows:

$$E_{circularity} = \frac{4\pi \times A}{P \times P}$$

where A refers to the area and P for perimeter of bounding box of each cell.



(a) (b) (c)

(d) (e) (f)

Fig. 7. Illustration of signal showing a start of cell division. (a)-(c) The *DIC* image frame 008.tif, 017.tif, 026.tif with black bounding boxes, respectively. (d)-(f) The *DIC* image frame 009.tif, 018.tif, 027.tif with indications on the start of mitosis as white bounding boxes, respectively.

Fig. 7 demonstrates a signal showing the start of cell division. (a)-(c) are three images at frames 8, 17, and 26, respectively in sequence 1 of *DIC-C2DH-HeLa* datasets. (d)-(f) correspond to images of (a)-(c) in the next frame namely 9, 18, and 27. White rectangles show successful detections in terms of a start of cell mitosis by the mean of the factors mentioned above.

D. Tracking

Because of the large number of cells in some frames, the Because of the large number of cells in some frames, the robustness and reliability were not powerful to use single distance feature to track cells in consecutive frames. To obtain a generalized automatic tracking module, the model needs to find some robust parameters that will not alter too much among different datasets in order to fit different motion modes and accommodate various topological features in the migration.

To handle the problem of tracking, the main feature that determines the identification of certain cells is the nearest-neighbor motion [15]. Nevertheless, through the experiment, it is clear that the movement of the cells also results in shape change and intensity changes, sometimes mitosis will also affect the judgment. Thus, a combination of features from several aspects is necessary to increase the robustness [16].

After obtaining the labels for cell segmentation, bounding rectangle and inscribed ellipse of each label were found. The Euclidean distances of two cells were computed as follows:

$$\mathrm{E}_{displacement} = \frac{\left\|\overrightarrow{x_{a1}^{t-1} x_{a2}^{t}}\right\|}{\sqrt{a^2 + b^2}}$$

where x_{a1}^{t-1} and x_{a2}^{t} represent the coordinates of centers of the bounding boxes for two cells in two consecutive frames, and a and b indicate the size (height and weight) of the frame image.

Then, as mentioned in [16], the coverage of two cells also implies the correspondence of two cells. The feature considering the ratio of overlapping area to original area was calculated by:

TABLE I
Cost Matrix for Cell Correspondence

Cell ID	a	b	c	d	e	…
A	0.048	0.426	0.559	0.617	0.500	…
B	0.414	0.073	0.540	0.617	0.503	…
C	0.548	0.537	0.038	0.490	0.630	…
D	0.456	0.447	0.576	0.661	0.458	…
E	0.611	0.615	0.487	0.087	0.682	…
…	…	…	…	…	…	…

$$E_{coverage} = 1 - \frac{\left(S_{overlap}^{t-1,t}\right)^2}{S_{a1}^{t-1} \times S_{a2}^{t}}$$

where $S_{overlap}^{t-1,t}$ refers to the overlapping area of two cells in consecutive frames and S_{a1}^{t-1} and S_{a2}^{t} indicates those two cells area alone.

Besides, the eccentricity of the deformation was also computed by fitting ellipses on the segmented cells and calculate the degree of deformation as follows:

$$E_{deformation} = \frac{|C_{a1}^{t-1} - C_{a2}^{t}|}{\sqrt{(C_{a1}^{t-1})^2 + (C_{a2}^{t})^2}}$$

$$C = \frac{P^2}{4\pi \times A^2}$$

where P refers to perimeter of the ellipses and A refers to the area, C_{a1}^{t-1} and C_{a2}^{t} represent the eccentricity of two cells in consecutive frames.

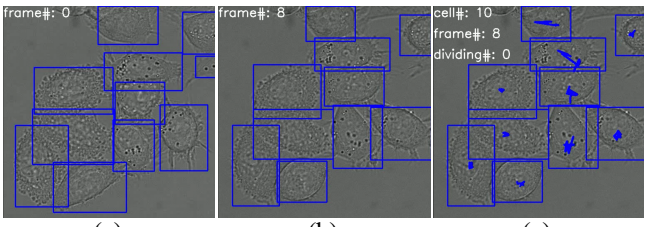
In this method, the cost matrix was applied to find the best correspondence between two cells in two consecutive frames. Normally by taking the minimal cost, we can get the correspondence for each cell, but due to the segmentation error, it is necessary to set a threshold α for the minimal error to avoid the incorrect connections to some extent.

In Table I, each cell had its corresponding cell in the previous frame unless the cost is above the threshold. For each cell in the current frame, it is feasible to trace back to its corresponding cell in the previous frame and all back to the position where it first appeared in the sequence. By doing so, the trajectory for each cell could be plotted as shown in Fig. 8.

E. Motion Analysis

Motion analysis is carried out in four angles: speed, total distance, net distance, and confinement ratio based on the trajectory history of a cell generated by Section III D. The speed is calculated as the movement per frame, the movement is estimated by the Euclidean distance between the centre point of the last frame and current frame. The total distance is formulated as the accumulative distance from the first frame to the current frame:

$$D_{total} = \sum_{i=1}^{k} speed_i$$



(a) (b) (c)

Fig. 8. Illustration of cell motion trajectory. (a) Image of *DIC-C2DH-HeLa* at frame 0 with bounding boxes. (b) Image of *DIC-C2DH-HeLa* at frame 8 with bounding boxes. (c) Image containing the trajectory of each cell from (a) to (b).

where k is the frame number of the current frame. The net distance is measured as the Euclidean distance between the centre point of the first position and current position. And lastly, the confinement ratio equals to total distance divided by net distance.

IV. Experimental Setup

The experimental codes were running on a macOS Mojave version 10.14.6, the code was written in Python 3.7.4 and tested on Jupyter Notebook platform.

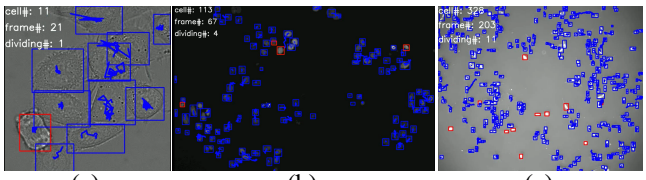
The cell segmentation method various between different datasets. For *DIC* dataset, explained detailly in Section V, are segmented using Deep Learning Techniques. In the DL model, in order to segment accurately, the cell mask and marker threshold are set to 230 and 240, respectively. The erosion size of *DIC-CDH-HeLa* in the morphological opening operation is 8. The boundary value is set to 15. This means the cell is removed when the cell distance is less than 15 pixels.

Regarding the augmentations process for the TTA, we use the same 7 augmentations as mentioned in Section III along with the original raw image: 1. Raw image 2. Horizontal flip 3. Vertical flip 4. Transpose 5. R(90°) 6. R(-45°) 7. R(45°) 8. Integration of transpose and R(90°).

The fluorescence microscopy is processed with parameters as follow: $l_b = 5$ to reduce noise whose size is smaller than 5×5, $l_w = 21$ to remove shading area larger than 21×21,$t = 0$ and $t, \tau = 100$ to remove the mask and segmented label smaller than 100 pixels. The phase-contrast microscopy requires similar parameters only with these differences: t =*Otsu's value* adapts to each frame's intensity, $t, \tau = 20$ and *blackHat=False* since we do not want to apply black top-hat filter on phase-contrast images.

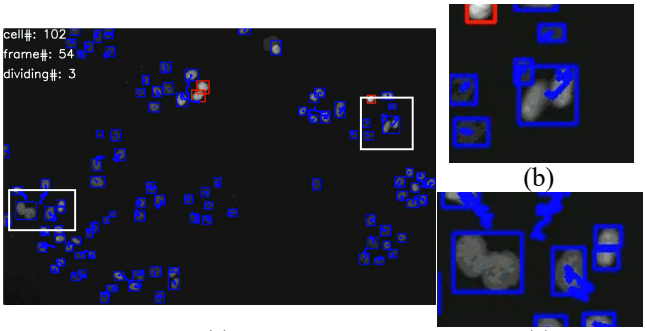
The tracking threshold parameter α is configured to a quantile of 0.9 for all the minimum error for each cell in the specific frame for dataset *Fluo* and *PhC* to remove the lower 10% associations in the cost adjacency matrix, since we are typically less confident with these associations, and is set directly 0.5 for *DIC* due to the small number of cell and high accuracy of segmentation. Besides, the weights for the cost matrix are 0.5 for displacement error, 0.4 for coverage error, and 0.1 for deformation cost.

In the mitosis detection module, these parameters were specially designed: $E_{intensity} = 123, 133, 236$ for *DIC*, *Fluo*, and *PhC*, separately, and $E_{circularity} = 0.9$ for all three datasets.



(a) (b) (c)

Fig. 9. Illustration of cell detection, motion trajectory, and mitosis detection. (a) Image of *DIC-C2DH-Hela* at time 21. (b) Image of *Fluo-N2DL-HeLa* at time. (c) Image of *PhC-C2DL-PSC* at time 203.



(a) (b) (c)

Fig. 10. Demonstration of the frame 54 in *Fluo*. (a) entire frame 54. (b) The enlarged detail of the segmentation error at right white box. (c) The enlarged detail of the detection error at left white box.

The output of the proposed method is evaluated against ground truth which is acquired by manual inspection on randomly sampled frames. The evaluation method is designed in three main parts.

The segmentation evaluation is assessed by two methods for accurate comparison. First, we calculate the nominal accuracy as the number of cells detected by the proposed method divided by the manual counted cell number. Second, we calculate the comparative accuracy as the manual counted the number of cells detected by the proposed method divided by the manual ground truth.

The mitosis detection evaluation calculates the precision and recall of which the cells division happens within certain frames when that cell is identified as in the process of mitosis. The number of frame is bounded by the number of sequential sampled frames. The tracking evaluation is measured as the correct association between the previous and current frames over the total number of associations.

V. Results and Discussion

A. Datasets

To demonstrate our method performance, we chose datasets: *DIC-C2DH-HeLa*, *Fluo-N2DL-HeLa*, *PhC-C2DL-PSC*. (see Fig. 6, Fig. 2, and Fig. 3 for examples respectively) from the international Cell Tracking Challenge (CTC) by the Computer Vision Course COMP9517 of The University of New South Wales. Datasets differ in cell size, cell appearance, and illumination condition.

The first dataset of HeLa cells *DIC-C2DH-Hela* was captured in 2010 by Gert van Cappellen from Erasmus Medical Center, Rotterdam, by differential interference contrast microscopy. There are 84 frames in each training sequence and 115 frames in each testing sequence. Only 9 training images in each sequence have masks. Each frame is represented by a grayscale image of size 512 × 512 pixels.

TABLE II
Evaluation of Cell Segmentation

Data	Ground truth	Automated count	Manual count	Accuracy (automated)	Accuracy (manual)
DIC	1191	1165	1165	97.82%	97.82%
Fluo	2325	2101	1943	90.37%	83.57%
PhC	9283	8927	8562	96.17%	92.23%

* The ground truth are manually inspected over randomly sample frames of entire sequence for *Fluo* and *PhC*. 28 sampled frames for *Fluo* and 40 sampled frames for *PhC*.

The second dataset, *Fluo-N2DL-HeLa*, was HeLa cells stably expressing *H2b-GFP* which captured by Mitocheck Consortium, by fluorescence microscope. There are 92 frames in each training sequence and 92 frames in each testing sequence. 28 masks and 8 masks in *Fluo-N2DL-HeLa* Sequence 1 and Sequence 2, respectively. One frame is represented by a grayscale image of size 1100 × 700 pixels.

The third dataset of pancreatic stem cells *PhC-C2DL-PSC* was captured by Dr. T. Becker from Fraunhofer Institution for Marine Biotechnology, Lübeck, Germany, in 2011. In each training sequence, there are 426 frames of size 720 × 576 pixels. Only 4 training images have full reference annotation.

Fig. 9 shows three processed results with cells segmented in blue bounding boxes, motion tracked in blue lines and mitosis labelled in red bounding boxes.

B. Evaluation of Segmentation

There are two methods to calculate the accuracy of cell detection as mentioned in Section IV, and the second method, the accuracy with manual counting of the segmented cells by the proposed method, is more representative to the actual accuracy. Take the example of frame 54 for *Fluo* in Fig. 10. The manual inspection of the number of cells in frame 54 is indeed 102, but clearly, there are some over-segmentation and under-segmentation that happened at the same frame shown in Fig. 10(b)(c). If the first accuracy approach was adopted, we would obtain 100% accuracy for this particular frame, while the second accuracy approach would discard all the segmentation error and gives $(102 - 7)/102 = 93\%$ where 7 is obtained by manual inspection over each bounding box. Meanwhile, even though the manual counting approach is closer to the real accuracy, it is more subjective as the number of cells increasing in the frame, and the automated approach would have a standardized accuracy. The overall accuracies of all three datasets are listed in TABLE II. The DIC dataset segmented by the U-Net approach is much more accurate due to its sophistication. The over-segmentation and under-segmentation were never an issue to this approach, but became unavoidable to the traditional watershed method. The *Fluo* dataset has less dynamic intensity range compares to the *PhC* dataset, it is harder to separate the foreground from the background, and therefore, has lower accuracy compares to the *PhC* dataset using the proposed method.

TABLE III
PERFORMANCE EVALUATION OF MITOSIS DETECTION

Data	TP	FP	FN	P	R	F-measure
DIC	8	0	1	100%	88.89%	94.12%
Fluo	14	9	10	60.87%	58.33%	59.57%
Phc	63	63	30	50%	67.74%	57.53%

* Similar sampling strategy was adopted as in TABLE II.

TABLE IV
ACCURACY OF CELL TRACKING

Data	Ground truth	Correct correspondence	Accuracy
DIC	385	298	77.40%
Fluo	2325	1897	81.59%
PhC	9283	8318	89.60%

* Similar sampling strategy was adopted as in TABLE II.

C. *Evaluation of Mitosis Detection*

Since the proposed method is unable to capture all mitosis events, it is better to evaluate the performance of mitosis detection with the parameters true positive (TP), false negative (FN), false positive (FP), precision (P), recall (R) and F-measure. The precision characterizes the correctness of the result and the recall characterizes the completeness of the result. The parameter P is calculated as the percentage of actual mitosis over detected mitosis, i.e., $TP/(TP+FP)$. The parameter R is obtained by the percentage of the detected true mitosis over the total mitosis, i.e., $TP/(TP+FN)$. The F-measurement is the harmonic mean of the P and R and is formulated as: $2\times(R\times P)/(R+P)$. It can be observed in TABLE III that the mitosis detection has higher performance for the *DIC* dataset while lower performance for *Fluo* and *PhC* dataset. This could due to the approaches in the cell division detection. The intensity and circularity features well capture the mitosis in *DIC*, and relatively poor demonstrated in *Fluo* and *PhC*. The accuracy of mitosis detection is also influenced by the actual dividing duration in term of frames. For example, the *Fluo* sequence lasts for 30 mins in 372 frames while PhC sequence lasts for 10mins in 1456 frames. Same number of frames limit employed on the mitosis duration of two cells sequence would expect lower accuracy on *PhC* than *Fluo*.

The performance of mitosis later affects the performance of cell tracking.

D. *Evaluation of Tracking*

The performance of tracking a single cell from its first frame to the last frame is constituted by many successful tracking of two consecutive frames, and therefore we evaluate the cell tracking by evaluating the accuracy rate of each cell correspondence between two frames. The accuracies of cell tracking on three datasets are shown in TABLE IV. Tracking performance on the *Fluo* dataset is mildly worse than *PhC*. One reasonable observation is that cells in *Fluo* moves more irregular, and the distance feature contributes more accuracy in *PhC* rather than in *Fluo*. Cell division frequency is another issue. Cells in *Fluo* divide more frequently than cells in *PhC*, and the mitosis detection we proposed are not highly accurate to overcome this error. It is also worth mentioning that even though the mitosis detection method works much better on the DIC dataset, it still has the worst tracking accuracy among all three datasets. This is due to the fact that cells in *DIC* are close to each other, short distance movement will cause the distance feature badly identifies the same cells. The cells in *DIC* also overlap with each other which makes the overlapped area feature less efficient.

VI. CONCLUSION

In this paper, we have shown our method for semi-automatic cell tracking of the time-lapse microscopic videos. This model presents a good performance for cell segmentation and a moderate effect of cell tracking and mitosis detection. Due to the promising image preprocessing procedure such as top-hat filters and ultimate erosion with the constraints of the iteration times, the occurrence of multiple local maxima in one object is reduced. Therefore, the probability of over-segmentation/under-segmentation has been vastly decreased. This provides a relatively reliable base condition for cell tracking. When it comes to the tracking stage, the cost matrix gives reasonable evidence for the matching of cells among frames, and the combination of topological features and motion features further improve the robustness among various datasets. As for mitosis tracking, the intensity and circularity of the cells indicate the high possibility of cell division. However, there are still some issues yet to be solved. There is some possibility that the cell tracking will be affected if there is a new cell that is close to other existing cells when it enters the frame, and the approach taken for mitosis detection works better in *DIC* images than others. Therefore, further study needs to be done to increase the robustness and accuracy. To conclude, our proposed method provides a good start of cell tracking algorithm and could be applied in some related biological applications to solve the problem regarding cell migration.

VII. CONTRIBUTION OF GROUP MEMBERS

Chen: DL research, cell segmentation and test-time augmentation on *DIC*.

Li: Section III B, some parts of Section IV, Section V A, and demo slides.

Ren: Cell division on three datasets and Section III C of report.

Xie: traditional cell segmentation on *DIC*, *Fluo* and *PhC*, motion analysis, result videos with key information, Section III A.B.E, Section IV, Section V B.C.D of report.

Zhou: early stage research, cell tracking, abstract, Section I, Section II, Section III D, Section VI of report.

VIII. REFERENCES

[1] O. Debeir, P. Van Ham, R. Kiss and C. Decaestecker, "Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes", IEEE Transactions on Medical Imaging, vol. 24, no. 6, pp. 697-711, 2005.

[2] C. Zimmer, E. Labruyere, V. Meas-Yedid, N. Guillen and J. Olivo-Marin, "Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing", IEEE Transactions on Medical Imaging, vol. 21, no. 10, pp. 1212-1221, 2002.

[3] X. Yang, H. Li and X. Zhou, "Nuclei Segmentation Using Marker-Controlled Watershed, Tracking Using Mean-Shift, and Kalman Filter in Time-Lapse Microscopy", IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 53, no. 11, pp. 2405-2414, 2006.

[4] K. Li, E. Miller, M. Chen, T. Kanade, L. Weiss and P. Campbell, "Cell population tracking and lineage construction with spatiotemporal context", Medical Image Analysis, vol. 12, no. 5, pp. 546-566, 2008.

[5] E. Meijering, O. Dzyubachyk, I. Smal, W. A. van Cappellen. Tracking in cell and developmental biology. *Seminars in Cell and Developmental Biology*, vol. 20, no. 8, pp. 894-902, October 2009.

[6] A. Sacan, H. Ferhatosmanoglu and H. Coskun, "CellTrack: an open-source software for cell tracking and motility analysis", *Bioinformatics*, vol. 24, no. 14, pp. 1647-1649, 2008.

[7] M. Möller, M. Burger, P. Dieterich and A. Schwab, "A framework for automated cell tracking in phase contrast microscopic videos based on normal velocities", *Journal of Visual Communication and Image Representation*, vol. 25, no. 2, pp. 396-409, 2014.

[8] N. Otsu, "Athreshold selectionmethod," IEEE Trans. Syst.,Man Cybern., vol. 9, no. 1, pp. 62–66, Jan. 1979.

[9] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," IEEE Trans Pattern Anal. Mach. Intell., vol. 13, no. 6, pp. 583–598, Jun. 1991.

[10] A. Chambolle, An algorithm for total variation minimization and applications, Journal of Mathematical Imaging and Vision, Springer, 2004, 20, 89-97.

[11] T. Sund & A. Møystad: Sliding window adaptive histogram equalization of intra-oral radiographs: effect on diagnostic quality. Dentomaxillofac Radiol. 2006 May;35(3):133-8.

[12] Meijering, E., Jacob, M., Sarria, J. C., Steiner, P., Hirling, H., Unser, M. (2004). Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. Cytometry Part A, 58(2), 167-176. DOI:10.1002/cyto.a.20022

[13] N. Moshkov, B. Mathe, A. Kertesz-Farkas, et al. Test-time augmentation for deep learning-based cell segmentation on microscopy images. Sci Rep 10, 5068 (2020). https://doi.org/10.1038/s41598-020-61808-3

[14] Lux, Filip & Matula, Petr. (2019). DIC Image Segmentation of Dense Cell Populations by Combining Deep Learning and Watershed. 236-239. 10.1109/ISBI.2019.8759594.

[15] C. Andrey and L. Andrey, "Tracking feature points: Dynamic programming algorithm", 2009 IEEE Congress on Evolutionary Computation, 2009.

[16] M. Dewan, M. Ahmad and M. Swamy, "Tracking Biological Cells in Time-Lapse Microscopy: An Adaptive Technique Combining Motion and Topological Features", IEEE Transactions on Biomedical Engineering, vol. 58, no. 6, pp. 1637-1647, 2011.