Cell tracking in time lapse microscopy imaging

John Cawood

Christopher Low

Sankalp Shukla

Alek Petkovski

***Abstract*— Cell biology is the examination of the 'building blocks' of life. Over the past few decades, time lapse microscopy has attracted the likes of computer scientists due to the need for automating cell tracking. The image sequences resulting from the microscopy imaging can contain large numbers of cells or there may exist too many image sequences to analyse by manual methods. There exist significant advantages in letting algorithms automate the process such as removing the need for the labour-intensive operation of counting and tracking the cells in addition to reducing human error. This report details a solution for cell segmentation and cell tracking and then further extending the program to detect mitotic events and track a single cell over a period of time through a combination of convolutional neural networks and several computer vision techniques. The results show varying accuracies between the cell datasets.**

I. INTRODUCTION

Cell biology is a field that observes and analyses the behaviours of living cells at the microscopic level. Through research, a better understanding of the 'building blocks of life' can be achieved. One such field that greatly benefits from examining cells is the field of medicine which can branch from observing cancer cell behaviour to drug administration.

Cells has seen a significant rise in interest in the past few decades due to improved microscopy capture techniques. Time lapse microscopy was developed to allow the observation of the cell activity over time through captured frames. Biologists are generally interested in how the cells move, undergo division and interact with their surroundings.

Computer vision scientists are tasked in developing code that help streamline the extensively laborious activity of manually counting hundreds to numerous thousands of cells in a single image. There currently exist many well-established computer vision techniques and programs that help cell tracking. However, ultimately, it is best to observe the imaging in question and develop an algorithm on a case by case basis.

This report looks at the international Cell Tracking Challenge (CTC) and how a group of 4 computer vision students develop code for detecting and tracking cells, detecting cell division and analysing the motion of a selected cell in the given sequences of images. The given images contain varying cell counts, sizes and shapes from different biology experiments. Cell segmentation is the process of detecting cells in an image. A well-established segmentation method includes 'watershed transform'[1]. This however works best for relatively low cell counts and images that don't exhibit excessive noise. In these cases, a combination of various computer vision processing techniques and the aid of convolutional neural networks can provide a solution. A run-through established computer vision techniques and how they assist in cell tracking is described. Our findings, discussion and potential future work of these results are presented at the end of the report.

II. LITERATURE REVIEW

This section details past techniques used in cell tracking which incorporate the segmentation of cells and tracking of cells over time and tracking the cell phenomenon known as mitosis.

Segmentation of cells is necessary in establishing where the cells are and creating outlines of the cells per image. References [1], [2], [3] and [4] outlines various methods used in cell segmentation as well as cell tracking. Segmentation is generally established by modelling bright areas against a dark background. However, there exists complex imagery which may contain for example substantial noise or a high cell count that simply comparing the foreground with the background won't be sufficient. These cases require specialised techniques or further processing. One established and popular solution used is the 'watershed transform' method. This method first forms a topography of the image and gradually 'fills' the valleys with water i.e. the lowest heights of the topography. As this water level rises, the image begins to distinguish regions of interest (cells) and the background. Reference [5] uses an algorithm based on this method to segment microscopic nuclei clusters with 90% of their test clusters of blood and bone marrow correctly segmented. This method however significantly falls short when the image contains noise and the possibility of segmenting excessively known as 'over segmentation'. Despite these drawbacks, pre-processing and post-processing have been developed to complement the watershed transformation to resolve these issues. Reference [6] and [7] makes use of top-hat, gaussian, erosion and dilation filtering as a form of processing to help watershed segmentation with high accuracy.

Reference [8] perform pre-processing to better enhance the images of cell colonies with use of Contrast-Limited Adaptive Histogram Equalization (CLAHE) in conjunction with other techniques. This resulted in higher definition of cell regions.

Another segmentation technique is the basic method of comparing each individual pixel of the image to a threshold value and through this establish the foreground and background. The simplicity of this method allows it to be implemented in cases where image complexity is low such that there exists high contrast between the cells and background as mentioned in reference [9]. Any further

complexity such as noise and autofluorescence (the natural emission of light from some cells) in the images would easily introduce errors. Reference [7] develops an image with the use of an adaptive threshold-based algorithm to automate the process of segmentation of calcium imaging. With a signal to noise ratio of less than 20%, the algorithm was able to produce higher than 80% precision. This algorithm works by iteratively calculating the threshold value.

Reference [10] mentions another method involving the use of templates of existing cell forms to outline the cells however due to the nature of cell behaviour, it doesn't work consistently when the cells begin to undergo mitosis or similar phenomenon.

A technique incorporating Convolutional Neural Networks (CNN) outlined in reference [11] works to automate cell classification of several human cell types. This method resulted in the algorithm performing better then classifying the cells then a human expert in actin-labelled fluorescence microscopy images. (97.6% vs 78.6%)

A relatively new method in cell segmentation involves the use of deformable models. These models allow the capturing of cell division as done so in reference [12]. By using a predetermined energy function and iteratively working through an image to solve the deformable model, one can combine the biological knowledge with the image properties. Unlike 'watershed transform' this method tends to under segment the cells especially when cells are adjacent to each other or are in contact. This however, like 'watershed transform' can be remedied through post processing.

The next step in tracking cells after segmentation is the association of cells between each subsequent frame. The importance of ensuring the cells are matched between images cannot be overstated when examining biological phenomenon. A simple approach would be to assume the cells move little or not at all in a spatial aspect such that a cell's centroid in an image would be the on or near the same location on the next image. A successful example of this is demonstrated in reference [13] where an tracking algorithm with Kalman Filter implemented used the centroids of cells. This proved to be highly successful compared to manual practices. This however starts to fall short when examining more complex biological cell phenomenon or if a there is a high cell count. However, by also using other features of a particular cell, the accuracy of linking the cells between frames can be improved. Such features include the cell's perimeter and area/volume.

Another established method of cell association is by extending some of the segmentation methods. The use of templates in cell segmentation can also be used in recording the features of a particular cell in one image and using this information to find the same cell in the next image. Reference [14] makes good use of this method to determine the movements of the cell in addition to its rotation. Deformable models too use a similar technique in which by segmenting one image, it can then be used to initialise the next image to help associate cells that share a similar location. Again, nevertheless this method doesn't work well when there are a high number of cells as it may incorrectly associate to other adjacent cells. Beyond these techniques exist more sophisticated methods which help to alleviate these issues such as gradient-vector flows, estimated cell dynamics, probabilistic schemes, and other unique processes to determine cell contact, appearances, and disappearances.

Reference [2] mentions that cell trajectory can be established through the use of linear interpolation of all the measured coordinates from previous images of the same cell. From this, several useful parameters can be determined to help aid the examination of biological processes such as the distance a cell had travelled over a set period of time, average/instantaneous velocity and directional change.

In reference [15], several features between each image are measured and used to track mitotic events. The features they used in particular are, a fitting error ratio, distance and variance of intensity between the closest component cells. With these features, the algorithm could detect mitosis despite overlapping cells and varying intensities. Reference [2] and [4] indicate there exists several programs that can detect cell division such as StarryNite, uTrack, CellProfiler and Trackmate. Furthermore, deep learning has been found to be useful in detecting mitosis as well as in conjunction with computer vision techniques to help segmentation. [4]

III. METHODS

This section details the techniques used for each of the tasks and why it was used as a solution to cell tracking for each of the given datasets over other techniques that were tested.

A. Segmentation of Dataset DIC-C2DH-HeLa

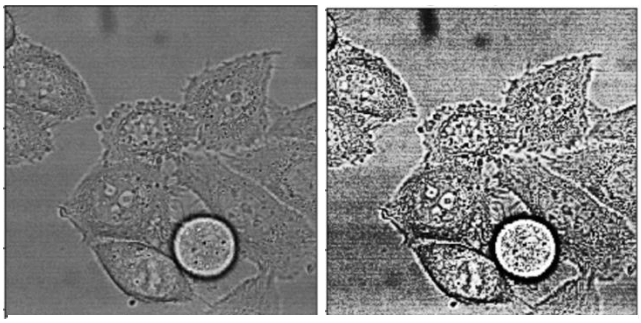
The segmentation of the DIC-C2DH-HeLa dataset proved to be a challenge due to the differing nature of this dataset from the others. The cell size of the dataset is a lot larger, and the nucleus of the cell is surrounded by noisy elements that make it difficult for many segmentation techniques to succeed. The low contrast between the background and the foreground elements of the image makes it even more of a challenge to properly identify the cell boundaries and separate the cells from each other. By taking into account all of these problems, an acceptable solution needed to be found to track the different cells in the given image sequences.

The low contrast image sequence first needs to be contrast equalized so that the cells can be distinct from the background of the image. A CLAHE (Contrast Limited Adaptive Histogram Equalization) function was used but it was later found to be underwhelming as the Histogram Equalization function inside the cv2 Python library proved to be more useful in differentiating the background from the foreground. Figure 1 shows the difference between the two equalization functions is shown. It is apparent from this figure that Histogram Equalization performs better than the CLAHE function.

Now, the noise inside the cells needed to be removed so that the cells can be properly separated. To achieve this, an image blurring function will be used. Hence, a gaussian filter function was used but it did not yield the expected results as the blurring either did not fully remove the noise or removed the boundaries of the cell along with the noise. A different blurring function called the median blur was then used because it was able to segment the cell boundaries and

reduced the noise inside the cells better than a gaussian blur filter.

The image now needed to be converted into a binary



(a) (b)

Figure 1 (a) CLAHE function applied to an image in sequence 3 of the DIC-C2DH-HeLa dataset, (b) Histogram equalization applied to an image in sequence 3 of the DIC-C2DH-HeLA dataset

image that differentiated between the cells and the background. A binary threshold was applied to the image, but it failed to work due to similarities between the cytoplasm of the cell and the background. H-dome filter was also used to obtain the regional maxima and further equalize the image for eventual segmentation. The results of using the binary filter before and after applying h-dome filter are seen in Figure 2. Figure 2a is unable to separate the cytoplasm of the cell from the background while Figure 2b is unable to distinctly define the edges of the cells. Both the images cannot be used for further segmentation of cells.

In order to improve this result different methods will need to be implemented. The use of binary threshold and the h-dome filter was discarded, and an adaptive threshold was used. This threshold yielded much better results and hence, was used for the final segmentation process.

During the process of attempting different methods, it was realized that two separate segmented cell images will need to be obtained so that one image can separate all the cells from the background and the other image can separate the cells from each other. Making use of this method of separation instead of doing both the processes using the same cv2 functions will make the problem easier to solve. After attempting different masking methods that yielded no satisfactory result, a u-net architecture was used. The given images and labels that amounted to a total of 18 images were used to train this network. This architecture provided us with more reliable image masks and hence, was used for the final segmentation of cells.

After the binary masked image was obtained from the u-net, it needed to be cleaned to remove artifacts. Since, the cell size was big anything lower than a certain threshold was removed using the 'remove_small_objects' function in the morphology package of cv2. Same function was applied to the image that was obtained from the cv2 functions. This is discussed in more detail in Section V Results and Discussion.

The final segmented image that was obtained was a combined image of the mask that was obtained from the u-net architecture and the binary image that was obtained from the cv2 functions. The two different image separated different parts of the image and were combined to give the final result. This method also provided an improved reliability in the final

(a) (b)

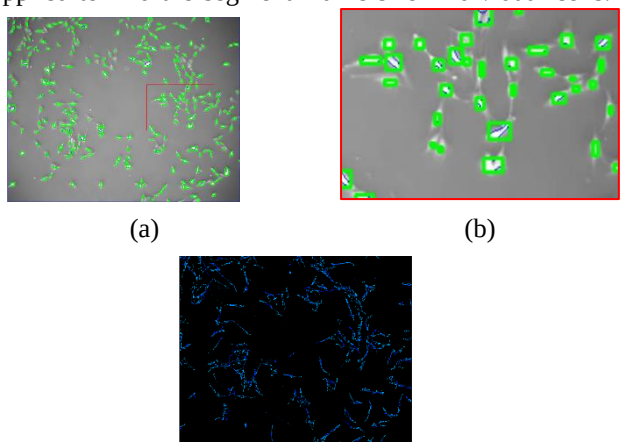
Figure 2 (a) Binary image before h-dome threshold is applied, (b) Binary image after h-dome filter is applied.

image as both the methods had to classify a cell as a cell in order for it to appear in the final image.

A. Segmentation of Dataset: Fluo-N2DL-HeLa and PhC-C2DL-PSC

For the Fluo-N2DL-HeLa Dataset the main issue is the extremely low contrast between the cells and the background. With the difference of intensity being less than 5 on the RGB scale between cells and the background. Because of this we made the assumption that for the cells to be considered for segmentation they need to be visible to an unimpaired human eye, since if this is not the case it is difficult to find distinction between artifacts, the background, and cells. If this assumption does not hold the failure should be attributed to the microscopy methodology and not the computer vision techniques used. Cells that are partially in frame on the edges of an image are also discounted as cells to be tracked.

To address the low contrast a gaussian blur is used to smooth an image with a square kernel size K=3. After this the blurred image is threshold using an adaptive Otsu technique [2]. To remove left over artifacts the erosion and dilation is used to perform opening and then closing on the image. An elliptical kernel of size K=5 is used for the artifact removal. After these pre-processing steps a watershed algorithm is applied to find the segment markers for individual cells. A



(a) (b)

(c)

Figure 4 (a) bounding boxes on an image in Fluo-N2DL-HeLa, (b) close up of subsect of (a), (c) centroids over time drawn to show cell pathways.

general form of pre-processing can thus be described as:

$$S = Closing(\, Opening(\, OtsuThresh(\, gaussianBlur(I, K = 3)), K' = 5), K' = 5)$$

Where S is the pre-processed image, I is the input image, K is a square kernel (size in pixels), and K' is an elliptical kernel.

The PhC-C2DL-PSC differs from the Fluo-N2DL-HeLa in that the contrast is not as low, the cell count is much higher, and the background is unevenly illuminated. Because of these differences the pre-processing methodology two additional steps are applied to the pre-processing methodology in the form of a top-hat filter for illumination correction and an extra erosion step for better separation in areas of tightly clustered and numerous cells. The difference can be described generally by altering the method for Fluo-N2DL-HeLa as follows:

$$I' = top - hat(I, K' = 15)$$

Where I' is a modified input to be passed into the prior method instead of I. As well as:

$$S' = erosion(S, K' = 5)$$

Where S' is a modified pre-processing result to be passed into a watershed segmentation as in the method for Fluo-N2DL-HeLa.

Bounding boxes are drawn around watershed markers to highlight individual cells and the centroid of each box is found and drawn at each timestep for tracking purposes. It is assumed that the centroid of a cell's bounding box approximates the centre of the cell itself.

B. Cell Tracking

The complexity of tracking biological cells increases with the number of individual cells to be analysed and the amount of movement imposed by each cell. This means to improve the accuracy of cell tracking, it's important to initially consider the datasets provided and account for sequences with a significant cell count. Tracking needs to consider the location of each cell from frame i to its location in frame i+1. Different methods for tracking were tested and observed to fail due to the complexity of the dataset. SIFT method was first tested in order to match features of cells from one frame to the next. Although it seemed to produce promising results for PhC-C2DL-PSC dataset due to the obvious variations in features, it failed to produce acceptable tracking results for Fluo-N2DL-HeLa and DIC-C2DH-HeLa due to the similarities of cell features. Another method this time for tracking based on features was incorporated. This method identified the basic features such as the area of the bounding box around the cell, the colour intensity of the cell and its ratio of height and width of the bounding box. The feature-based cell matching method also yielded inaccurate results with Fluo-N2DL-HeLa datasets but showed better results for DIC-C2DH-HeLa and PhC-C2DL-PSC.

A centroid based tracking approach was evident to produce acceptable results for all three datasets and hence why was deemed appropriate for the tracking of biological cells in this project. In almost all cases, each cell remains within the same spatial location to its consecutive frame meaning a comparison of centroids from different frames is crucial.

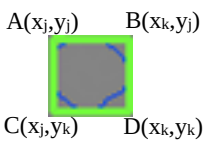


Figure 5 Cell centroid determined from bounding box.

A cells centroid (C) is defined as the centre co-ordinates of the cells bounding box. The corners of the box A, B, C and D are calculated, and the centroid is determined by Figure 5.

Cells are tracked by matching each cell in frame I to the cell in frame i+1 with the Shortest Euclidian Distance (SED).

SED of a cell is determined by calculating the Euclidian distance of the target cell from frame I to all the cells in frame i+1. The values are sorted to output the minimum value denoted as the cell with the minimum distance to the target cell.

Cell in frame i+1 = min (Euclidean distance (cell in frame i, all cells in frame i+1))

Each cell has its own unique identification value that is allocated when the cell first enters the system either as an initial cell in frame i or as a daughter cell of a new mitosis event. A cell in two consecutive frames is tracked when all the cells along with their identification values are matched.

C. Cell Mitosis

Task 2 involves detecting mitosis events which is the process of splitting one cell, denoted as the mother cell, into two daughter cells. The overall idea of detecting a mitosis event is to observe the cell in an arbitrary frame i and frame i+1 and make a judgement whether the cell has divided in frame i+2. During the stage of mitosis, a cells shape is changed either by becoming observably larger than its previous frame or getting deformed in the x and y plane. Therefore, if a cell in frame i+1 has an area that is significantly larger or small than an thresh value dependant on the dataset, that cell would be in a "pre-mitosis" state. Therefore, for a cell to be in the "pre-mitosis" stage:

"Pre-mitosis" stage when Area of Cell in frame i / Area of Cell in frame i+1 > thresh ratio.

A cell then undergoes mitosis, and the daughter cells are determined by evaluating the Shortest Euclidian Distance of the deceased mother cell's centroid to the centroids of all the new cells in the current frame. Calculating this will ensure that the relationship between mother cell and daughter cells will maintained. The daughter cells are classified as new cells and therefore have their own unique identification number assigned to them.

D. Analysis of Cell Motion

Cells move at various speeds and move around their environment colliding with other cells and proteins. In order

(a) (b)

Figure 6 (a) Before applying binary image cleansing techniques, (b) After applying image cleansing techniques.

to determine calculations of speed and distance, information of the cell's centroid must be stored.

Calculating the speed of the cells is found by taking the distance over the time. The distance is a measurement of the Euclidian distance between two centroid points of that cell. This means the distance is denoted as the centroid value of the cell in the current frame compared to the centroid value of that same cell in the previous frame. Time is denoted as the amount of frames occurred in the calculation.

Therefore, Speed of Cell = Euclidean distance (Centroid of Celli+1, Centroid of Celli) pixels/1 frame

The total distance that a cell travels is the sum of the Euclidean distances between each frame. This is:

$$Total\ Distance = \sum_{1}^{n} Euclidean\ Distance$$
$$(Centroid\ of\ Cell_i, Centroid\ of\ Cell_{i-1})$$

The net distance travelled by a particular cell is a calculation of the Euclidean distance of the current centroid state to the first instance where the cell is observed. This is shown as:

$$Net\ Distance = Euclidean\ Distance$$
$$(Centroid\ of\ Cell_i, Centroid\ of\ Cell_0)$$

The confinement ratio of cell motion is defined as the Total Distance travelled by a cell over the Net Distance travelled by the same cell.

$$Confinement\ Ratio = \frac{Total\ Distance}{Net\ Distance}$$

IV. EXPERIMENTAL SETUP

This section explains in greater detail the code developed in finding a solution to each of the tasks. With the chosen technique mentioned in the methods section, further detail in how the solution was approached and developed is explained for each of the tasks.

The final segmented image that was obtained from the DIC-C2DH-HeLa dataset made use of all the recommendations made in the Methods section of the report. In order to obtain the best possible result, the different parameters of the functions had to be optimized and the image manipulation had to be done in an optimal order.

First, a histogram equalization of applied to the image followed by a median blur and denoising function. A 5x5 rectangular kernel was used in all the functions. The median blur made the edges of the cell more prominent while also removing the noise from the cytoplasm. The denoising function removed any other artifacts that contributed to the noise in the image including the noise that was left over from the median blur function in the background and the foreground. The denoised image with the adaptive threshold applied can be seen in Figure 6a. Figure 6b shows the same image after binary image cleansing. The opening function followed by a closing function was applied to the image to remove any small artifacts from the image. Finally, a 'remove_small_objects' function was applied to remove artifacts that were not removed from the opening and closing functions. This was possible on this dataset as the cell size was fairly big and any segmented cell will be above a certain threshold. Different parameters were experimented and the parameter that provided the least number of false positives was chosen.

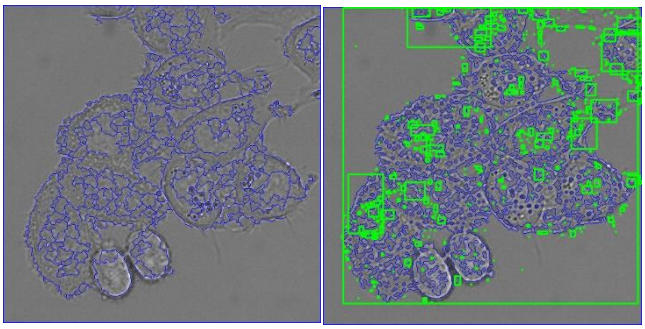
(a) (b)

Figure 7 (a) Image obtained from the u-net architecture, (b) Combined image obtained from images in Figure 7a and Figure 6b.

The u-net architecture was used to obtain the masks for the cells. The code was run on Google Colab and the segmented images were downloaded. The masking image had to be properly cleansed before it can be used as mask. Similar techniques were used to cleanse this image. An erode function was applied as the mask that was obtained from the algorithm stretched to the cell boundaries which introduced unnecessary artifacts. The iterations were set to 2 as this produced the most reliable result for the segmentation.

The last step of the segmentation process was to now combine the two binary images obtained from the u-net and cv2 functions to obtain a final image. This process of using the u-net as a mask on the cv2 image is shown in Figure 7. Figure 7a shows the image from the u-net algorithm whereas figure 7b shows the combined image obtained from the combination of images in Figure 7a and Figure 6b. It is apparent from the figure that the final image that is produced provides a good segmentation of cell and allows for tracking to be conducted on sequence 3 and 4 of this dataset.

Initially various pre-processing techniques to be passed into watershed segmentation were tested on each dataset such as top-hat and black-hat filtering, gaussian and median smoothing, different levels of opening and closing with various filter sizes and shapes, histogram equalization, subtraction of inverse thresholds and ridge detection. None of these methods gave adequate segmentation results on DIC-C2DH-HeLa. However, through testing these techniques a robust method of pre-processing was achieved for PhC-C2DL-PSC and Fluo-N2DL-HeLa as was described in the prior section.



(a) (b)

Figure 8 (a) over segmentation example of failed methods, (b) failure of the same method used for C2DL-PSC and Fluo-N2DL-HeLa on C2DH-HeLa.

To determine the quality of the segmentation average accuracy (of the cell count bounded) was calculated over a small subsect of sequence 1 for PhC-C2DL-PSC and Fluo-N2DL-HeLa. The reason this metric was used is because the cell count reaches over 300 cells in a timestep and thus more complicated methods would be very time intensive to determine manual ground truths for. Another consideration in choosing different methods was time efficiency, for PhC-C2DL-PSC and Fluo-N2DL-HeLa an intel® core™ i7-6700K CPU at 4.00GHz was used to run and evaluate the program.

The program itself was written in python and utilized the OPENCV library and its functions for pre-processing and watershed segmentation on PhC-C2DL-PSC and Fluo-N2DL-HeLa.

After the cells have been segmented from the images in each dataset resulting in the formation of bounding boxes, an implementation of tracking can be computed. Tracking utilizes the coordinated of the bounding box to calculate the centroid value of each cell. A tracking class is implemented for tracking and determining whether a cell will undergo mitosis. This class stores a different ordered dictionary for centroid values, coordinates, cells that are temporarily lost along with the corresponding cell id. The id of cells that are about to undergo mitosis is also stored in an array as well as the last cell id that has been used. An object is initialized at the start of the main function and with every new frame the objects parameters gets updated. This updates any new cells that have been detected in the segmentation process and also updates with the new coordinates and centroid values. In our class, SED of previous and current cell centroid values are determined in order to track cells from one frame to the next. Once the algorithm matches each cells, the new coordinates are appended to the location of all coordinated for that particular cell's ID. In this step, it's also determined whether the cell will go through mitosis by comparing the ratio of current area of the cell's bounding box to previous area value of the cells bounding box. If this value is larger than a thresh value of 1.6 (experimentally determined) then the cell's ID is appended to a mitosis array. In the next iteration the mitosis array is read and each mother cell is compared to the daughter cells within the proximity. The SED determines which daughter cells correspond to the mother cell. An output of this information is presented in the terminal as well as the count of the number of cells undergoing mitosis.

Motion of cells can be analysed as the centroid history of each values are stored. Using the formulas identified in the method section of this report, values can be deduced for the speed of the cell, total distance travelled, net distance travelled and also the confinement ratio. Pressing the 'p' button on the keyboard will output a user input box in the terminal where a cell to analyse can be chosen. With each frame the program will calculate the motion values and output them in the terminal for further analysis. Pressing the 'p' button again will change the target cell for analysis which results in different outputs in the terminal. When a cell no longer exists denoted by the id not visible on the output screen either due to the cell undergoing mitosis and producing two new daughter cells or simply no longer existing in the frame, then the centroid values are not being calculated for that cell and the analysis output will cease to exist.

```
########################################
Number of Cells in Image(t002.tif): 28

Predicted Mitosis 20:[120 529]
Predicted Mitosis 1:[678  15]

Mother Cell's ID is 3:(403, 173)
Daughter Cell 1 ID is 27:[426 151]
Daughter Cell 2 ID is 28:[402 183]

Count of Dividing Cells 2
########################################
Number of Cells in Image(t003.tif): 30

Predicted Mitosis 1:[689  30]

Mother Cell's ID is 20:(120, 529)
Daughter Cell 1 ID is 29:[124 550]
Daughter Cell 2 ID is 31:[132 514]

Mother Cell's ID is 1:(678, 15)
Daughter Cell 1 ID is 0:[637  21]
Daughter Cell 2 ID is 32:[689  30]

Count of Dividing Cells 1
```

Figure 9 Cell count and mitosis analysis

```
====================================
Speed of the cell  12 is  2.23606797749979  pixels/frame
Total Distance of the cell : 12 is  18.47603286813174
Net Distance of the cell : 12 is  13.152946437965905
Ratio of the cell motion: 12 is  1.4047067670556899
====================================
```

Figure 10 Cell properties

V. RESULTS AND DISCUSSION

This section details our results for each of the tasks according to each dataset followed by a discussion.

A. Evaluation of Segmentation

To display the effectiveness of the method used for the segmentation of the DIC-C2DH-HeLa dataset, the TP (True Positive), FN (False Negative), FP (False Positive) parameters were calculated. The results for sequence 3 can be seen in Table 1.

Table 1: Statistical parameters for sequence 3

Sequence	Ground Truth (Cells)	TP (Cells)	FP (Cells)	FN (Cells)
3	1468	1129	388	339

From this table, the value for the precision and recall can be calculated. The precision (P) and recall (R) are given by the formula below.

$$P = \frac{TP}{TP + FP} = \frac{1129}{1129 + 388} = 74.4\%$$

$$R = \frac{TP}{TP + FN} = \frac{1129}{1129 + 339} = 76.9\%$$

The similar precision and recall value show that the algorithm has no bias towards either the detection or non-detection of cells. This makes sense in the context of cell image analysis as there can be no inherent bias in cv2 functions or the initially labelled images. However, the relatively low value for both parameters show that the algorithm is not perfect and there were a lot of cases in the sequence where it failed. After individual analysis of the frames, it was determined that more often than not the algorithm did not detect the cells that were close to the boundary of the image. The parameters can hence be improved by eliminating boundary cells from the calculation as they do provide a full view of the cells. As more mitosis events occurred, the number of cells in the image sequence increased and the accuracy of the model decreased. The cells formed after the mitosis were detected more easily than the other cells as they had defined boundaries. Furthermore, when the cells occupied the whole image and the background size was minimal the algorithm was not able to separate the different cells from one another which decreased the number of cells detected and hence, increased the inaccuracy. Another reason for the inaccuracy of the method was the less defined cell boundaries as the image sequence progressed. This made it hard for the method to differentiate the individual cells. The cells in this image sequence did not move with a large velocity which helped in the identification of similar cells in consecutive images. This was one advantage of the DIC-C2DH-HeLa dataset over the other 2. To further improve the result of the segmentation the u-net will have to be trained with more labelled images which can provide a good base for the prediction of masks.

For PhC-C2DL-PSC and Fluo-N2DL-HeLa simple numerical accuracy was used to evaluate the segmentation performance.

Table 2 Average accuracy of cell counts for early and late in sequence.

Dataset	AL % (±5)	AH % (± 5)
Fluo-N2DL-HeLa	96.8	84.0
PhC-C2DL-PSC	99.6	96.1

In Table 2 AL denotes the average accuracy of a subsect of sequence 1 before cell count reaches 40 in Fluo-N2DL-HeLa and 100 in PhC-C2DL-PSC, AH denotes the average accuracy of the cell count with more than 80 cells in PhC-C2DL-PSC and more than 280 cells in Fluo-N2DL-HeLa.

From Table 2 it can be seen that the accuracy of segmentation drops off with higher cell counts especially in the Fluo-N2DL-HeLa dataset, this is due to extremely low contrast cell clusters that the program does not always detect.

Any given segmentation and centroid tracking step for Fluo-N2DL-HeLa and PhC-C2DL-PSC takes less than 0.3s to compute on average using the CPU mentioned previously. This is very good time efficiency for the degree of accuracy achieved and is the strong point of using this method above more complicated ones with higher accuracy, as they may display diminishing gains in terms of accuracy vs efficiency.

B. Evaluation of Tracking and Cell Mitosis

Detection of cells for dataset 1 proved to be harder than the detection of cells for dataset 2 and 3 due to the much lower accuracy of the segmentation. Even though the segmentation identified the cell locations, it could not properly identify the cell sizes. This made the detection of mitosis even tougher as the algorithm depended on calculating the size of the bounding boxes and then comparing that size to the previous frame in the image. The changing size of the bounding boxes and also the cells themselves resulted in inaccurate calculation for the centroid. Incorrect centroid detection causes the tracking technique to fail and resulted in incorrect identification of cell IDs. Furthermore, this increase in the size of the bounding boxes will result in a large number of false positive cases for mitosis detection.

The bounding boxes of cells in consecutive frames of dataset 1 were much smaller than the cells themselves due to the cell boundaries being ambiguous. The smaller bounding boxes not only did not detect cell movement accurately but also resulted in inaccurate detection of mitosis. When the full cell was finally segmented instead of just a small area of the cell in a consecutive frame the algorithm triggered a false positive which resulted in lower precision and recall scores and a high number of false positives. Table 3 shows the total number of mitosis events that occurred in dataset 1 and shows the TP, FP and FN parameters.

Table 3: Compares the mitosis events detected by the algorithm to the number of total mitosis events that occurred in the sequence.

Sequence	Ground Truth (Events)	TP (Events)	FP (Events)	FN (Events)
3	3	2	5	1

For better mitosis detection of dataset 1, the algorithm can be modified to identify perfect contour circles as it was noticed that the cells that underwent mitosis first formed an almost perfect circle and then divided into 2 cells. The cells undergoing mitosis also had darker boundaries which made it easier to detect them. This easier detection did not result in easier detection of the events due to the nature of the identification of algorithm and also the inaccuracies in the segmentation. Even though the mitosis cells were correctly identified, the algorithm designed worked better for dataset 2 and 3.

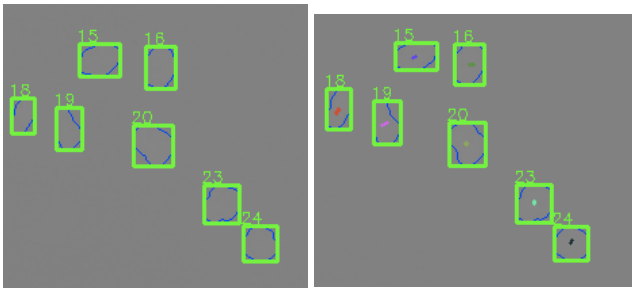
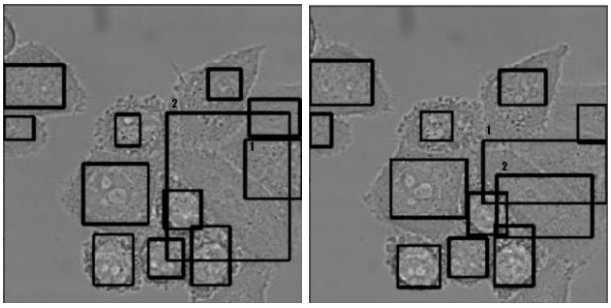


Figure 11 Results of cell tracking



(a) (b)

Figure 12 (a) Cell segmentation on image t016.tif showing bounding boxes, (b) Cell segmentation on image t017.tif showing the bounding boxes.

The accuracy of the results of cell motion are largely dependent on the accuracy of the segmentation and tracking of the datasets. For the DIC-C2DH-HeLa dataset, the precision and accuracy of the segmentation were 74.4% and 76.9% respectively. This inaccuracy would further lead into inaccurate results for the tracking of cell motion. As a result, the accuracy for the different motion parameters would also deteriorate. Furthermore, the cells in this dataset change their size without necessarily changing their position on the image. This results in different sized bounding boxes for the cells in consecutive images. Due to the usage of centroid tracking, the centre of the bound box changes position with the changing size of the box. This causes the speed parameter of the cell to be greater than the actual speed of the cell.

The inaccuracies in the DIC-C2DH-HeLa dataset range from false detection of cells in consecutive frames (FP) and no detection of cells (FN). These parameters are quantitative; however, the chosen method also results in qualitative inaccuracies such as the incorrect combination of cells that are close together. This means that, two bordering cells are combined to be one single cell in one image and then separated again in the next image. This would result in a huge change in the bounding box area and centroid detection which would in turn result in greatly increasing the speed of the cell if the cell was stationary. An example of this is shown in Figure 12 that analyses consecutive images in sequence 3 of the DIC-C2DH-HeLa dataset. Image in Figure 12(a) has a large bounding box 2 for the corresponding cell and a significantly smaller bounding box 1 for the other cell. The next image in the sequence in Figure 12(b) incorrectly combines the cell boundaries and bounding box 1 ends up being larger than bounding box 2. This shows that even in consecutive images, the speed of the cell will be highly inaccurate, and this problem will have to be remedied before this dataset can be segmented. Similar issues arose when calculating the distance and confinement ratios of the cells as they would also be affected by this kind of inaccurate segmentation.

A fail safe can be added to the already existing algorithm where a sudden increase in the bounding box of a cell would trigger a function which ignores a single image if the bounding box area is significantly increased and continues tracking that cell from the next image. Another algorithm will need to be implemented which does not falsely detect mitosis as one of these cases. This can be easily done as mitosis would result in a decrease in size of the bounding box and the other instance of inaccuracy would result in an increase and then decrease in the bounding box size.

Results for datasets 2 and 3 were more promising than the results for dataset 1 due to the high accuracy and proper bounding boxes. These results were shown in the previous section of the report. The accurate results for cell mitosis and tracking resulted in accurate values for the motion parameters. Additionally, the use of watershed for dataset 2 and 3 provided these accurate results for segmentation which resulted in accurate values. Watershed was not used for dataset 1 due to the masking image and the cv2 image producing inaccurate watershed markers which in reality amplified the inaccuracy instead of properly segmenting the cells. However, even small discrepancies in bounding boxes would still result in inaccurate values for the parameters. There were a few cases when the segmentation algorithm could not separate cells that were close together when the cell count was high. This was generally during the later stages of the image sequence. The inaccurate values in these later stages would only improve with proper segmentation even when the cell count and density is high. Furthermore, the segmentation sometimes incorrectly identified the size of the cell due to the nucleus being brighter than the cytoplasm. Again, this caused the centroid to move in the subsequent image, which resulted in inaccurate values for the speed, distance and confinement ratio parameters.

In order to get better predictions for the speed and distance of the cells for all the datasets, an analysis which uses more than just the last frame of the sequence will need to be done. A program that can make a loose prediction on the direction of cell movement and can also calculate the acceleration based on previous frames will be able to eliminate the discrepancies related to centroid tracking. This kind of program would be able to tell whether a sudden change in velocity is because of an inaccurate detection or an accurate prediction. For example, a cell moving in one direction would first have to show instances of deceleration which can then be attributed to a change in direction. A program like this can be designed and implemented in future works for better tracking.

The results for Fluo-N2DL-HeLa dataset portrayed very accurate results for tracking. The only limitation was when the segmentation process produced false positive bounding boxes due to inaccuracies with segmenting. If false positives were evident in a similar location between two consecutive frames, then the system may trace the bounding box as well as the other cells. This doesn't have a detrimental affect over the duration of the analysis of cells in the system, however, does affect the cell count produced in the current frame. This means further statistical analysis such as predicting the number of cells due to mitosis in an nth frame could be inaccurate as the false positive results would affect the model. However, overall, the tracking preformed to a high level of accuracy. Detection of mitosis events also proved to perform with high accuracy. This was due to the accurate segmentation of cells and tracking of cells from frame to frame. Most cells that divided were shown and labelled with the correct relationship of mother to daughter cell. The was evidence where the cell id would change of a cell, denoting that the cell has divided and that the result was the daughter cell when in fact no cell division was evident. This error occurred due to a poor consistent segmentation of cells size.

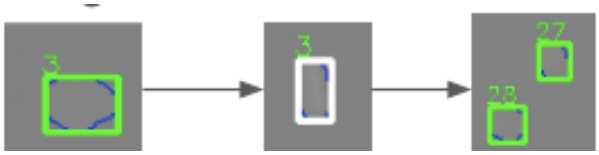


Figure 13 Mitosis results

This means that there were cases where the bounding box of the segmentation from
frame to frame would increase, resulting in an increase of area hence, achieving the mitosis threshold.

Since the segmentation accuracy for the PhC-C2DL-PSC dataset is above 95% even when there are above 250 cells in an image this results in precise centroid association between frames. False positive mitosis events are also reliant on the segmentation accuracy. With the highest accuracy in segmentation of all the datasets PhC-C2DL-PSC produces the best tracking results as well. The main source of error for mitosis detection is in cells that distort shape and result in inconsistent segmentation and centroid tracking between frames. This could be improved with a retroactive analysis of parent and child cells to correct inconsistent frame results and increase the accuracy of mitosis tracking.

VI. Conclusion

Cell biology is a science interested in the idiosyncrasies of life at the microscopic level from cell division to cell-cell interactions. This field has expanded to unprecedented levels due to improved imaging techniques in the past few decades which have attracted different fields of expertise beyond biology to analyse and help produce tangible results. One particular interest for computer scientists is cell tracking. Time lapse microscopy proved to be an effective method of capturing numerous images over any period of time for almost any cell biological processes. Traditionally, one would observe the cells frame by frame. Although some of these images can be observed and analysed by manual methods, others presented efficiency issues especially when dealing with very large numbers of images. By developing computer vision techniques, the analysis of microscopy images can be streamlined to requiring minimal input from human methods and allow significant progression in cell tracking.

Our solution is a mixture of several computer vision techniques and the use of convolutional neural networks to achieve cell segmentation and tracking including mitosis events. However, our results varied from dataset-to-dataset thanks due to varying numbers of cells, shapes of the cells, cell interactions and image noise. Furthermore, it is found that the accuracy of the tracking and detecting mitosis is heavily reliant on the cell segmentation accuracy.

Although the tasks given were completed, the program can be further refined to achieve more accurate results. A simple method is to acquire more data with different cell counts and image properties to further develop the program for different cases. Such data can train the convolutional neural network to better find the location of the cells between frames. Tracking cells can be improved by implementing a simple online and real time tracking algorithm (SORT) like the Kalman filter tracker.

VII. Contribution of Group Members

John Cawood developed code for segmentation, presented and written his methods, experimental setup, results, and discussion of datasets: Fluo-N2DL-HeLa and PhC-C2DL-PSC
Christopher Low performed the editing of the final report and written the introduction, literature review and conclusion. Furthermore, presented the future work and concluded the demo.
Alek Petkovski developed code for cell division, presented and written his methods, experimental setup, results, and discussion of all of the datasets.
Sankalp Shukla developed code for segmentation, presented and written his methods, experimental setup, results, and discussion of dataset: DIC-C2DH-HeLa as well as results and discussion for task 2 of dataset 1 and task 3.

Acknowledgment

Thanks to Justin Liang for introducing us to u-Net code. Thanks to the creator of u-Net code, ‘zhixuhao’. Thanks to Google for the use of Colab to run our u-Net.

VIII. References

[1] 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. https://doi.org/10.1016/j.semcdb.2009.07.004

[2] E. Meijering et al. Methods for cell and particle tracking. Methods in Enzymology, vol. 504, no. 9, pp. 183- 200, February 2012. https://doi.org/10.1016/B978-0-12-391857-4.00009-4

[3] V. Ulman et al. An objective comparison of cell-tracking algorithms. Nature Methods, vol. 14, no. 2, pp. 1141-1152, December 2017. https://doi.org/10.1038/nmeth.4473

[4] E. Moen et al. Deep learning for cellular image analysis. Nature Methods, vol. 16, no. 12, pp. 1233-1246, December 2019. https://doi.org/10.1038/s41592-019-0403-1

[5] Malpica, N. *et al.* Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry* **28**, 289–297 (1997).

[6] M. Ali Akber Dewan, M. Omair Ahmad, and M. N. S. 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. JUNE 2011*

[7] Shen, S. P., Tseng, H. A., Hansen, K. R., Wu, R., Gritton, H. J., Si, J., & Han, X. (2018). Automatic Cell Segmentation by Adaptive Thresholding (ACSAT) for Large-Scale Calcium Imaging Datasets. *eNeuro*, *5*(5), https://doi.org/10.1523/ENEURO.0056-18.2018

[8] Balomenos, A.D., Tsakanikas, P., Aspridou, Z. *et al.* Image analysis driven single-cell analytics for systems microbiology. *BMC Syst Biol* **11,** 43 (2017). https://doi.org/10.1186/s12918-017-0399-z

[9] Meijering E, Smal I, Dzyubachyk O, Olivo-Marin JC. Time-lapse imaging. In: Wu Q, Merchant FA, Castleman KR, editors. Microscope image processing. Burlington, MA: Elsevier Academic Press; 2008. p. 401–40.

[10] Kachouie NN, Fieguth P, Ramunas J, Jervis E. Probabilistic model-based cell tracking. International Journal of Biomedical Imaging 2006:1–10.

[11] Oei RW, Hou G, Liu F, Zhong J, Zhang J, et al. (2019) Convolutional neural network for cell classification using microscope images of intracellular actin networks. https://doi.org/10.1371/journal.pone.0213626

[12] Li K, Miller ED, Chen M, Kanade T, Weiss LE, Campbell PG. Cell popula- tion tracking and lineage construction with spatiotemporal context. Medical Image Analysis 2008;12:546–66.

[13] Huth, Johannes & Buchholz, Malte & Kraus, Johann & Schmucker, Martin & von Wichert, Goetz & Krndija, Denis & Seufferlein, Thomas & Gress, Thomas & Kestler, Hans. (2010). Significantly improved precision of cell migration analysis in time-lapse video microscopy through use of a fully automated tracking system. BMC cell biology. 11. 24. 10.1186/1471-2121-11-24.

[14] Wilson CA, Theriot JA. A correlation-based approach to calculate rotation and translation of moving cells. IEEE Transactions on Image Processing 2006;15:1939–51.

[15] Y. Li et al. Detection and tracking of overlapping cell nuclei for large scale mitosis analyses. BMC Bioinformatics, vol. 17, no. 1, p. 183, April 2016. https://doi.org/10.1186/s12859-016-1030-9

[16] ‘zhixuhao’, unet for image segmentaion. https://github.com/zhixuhao/unet