ON TRACKING SPINAL DISC CELLS

KP Dempsey[1], KP Lam[1], SJ Brown[2], JB Richardson[2]

[1]Computer Science and Mathematics, University of Keele, Staffordshire, UK
[2]The Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Foundation Trust, Oswestry, Shropshire, UK

ABSTRACT

This paper proposes the combined use of advanced texture based analytics and the Normalised Cross Correlation procedure to facilitate in vitro cells tracking with high-throughput time-lapsed video microscopy. The effectiveness of this technique was examined in the context of unlabelled phase contrast (PC) imaging, with the goal to investigate the life cycle of back disc cells. Computationally, the typically low contrast PC image adds to the challenges of cell tracking, which is often exacerbated by the highly varying cell morphology. To accurately determine the individual cell locations, the proposed algorithm computes the dynamic set of textural features for each cell, allowing its accurate identification in successive image/time frames. More importantly, such features also help calculate the location trajectory of every cell, for which further quantitative analysis can be developed, as proof of principle, to facilitate granulometric studies of vesicles present in these individually tracked cells. Preliminary results were compared to the manually constructed ground truth, allowing benchmarking of the proposed tracking algorithm in parallel with a freely available cell tracking platform.
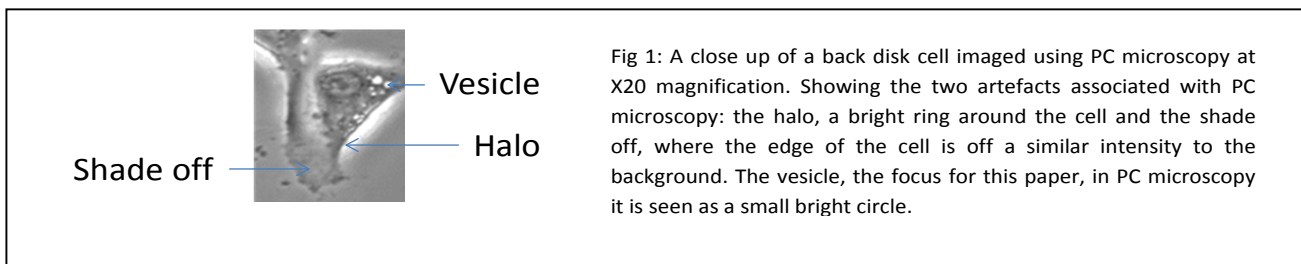
1. INTRODUCTION

In the past decade, cell/lineage tracking has started to become commercially and freely available. Further studies of the cell dynamics have led to a better understanding of the cell properties, since a detailed mapping between the dynamics of a cell and its behaviour can only be achieved with a complete understanding of the properties of the constituent parts of a cell. However, the majority of software developed to date has been designed to facilitate studying these intracellular parts of a cell with fluorescence labelling, as the latter is often used exclusively to help reveal detailed structures of the cell. Nevertheless, some constituent parts, such as vesicles and nucleoids, could be made visible under unlabelled PC microscopy. This paper presents a novel algorithm which enables automated tracking of these cells captured in time lapsed image/video sequences, with a specific focus on the detection of vesicles, over a three-day time period.

In cell biology, a vesicle is a small organelle within a cell, consisting of fluid enclosed by a lipid bilayer membrane. Vesicles perform a variety of vital cell functions including metabolism, transport, buoyancy control, and enzyme storage [1]. As such, they are particularly important for the trafficking of substances into and out of cells. For example, *lysosomes* are vesicles with enzymes which carry out functions to break down worn-out organelles, waste materials and substances/molecules that the cell may engulf. Thus, the absence of vesicles within a cell could lead to a build-up of the waste substances that could undermine the proper functioning of the cell. Given this large array of functionalities, it should be apparent that the study of vesicles is becoming increasingly important in all fields of cell biology. At present, however, whilst there are many cell tools exist to track cells, few, if any, is capable of monitoring the vesicles inside the individual cell particularly when it is unlabelled.

Broadly speaking, previous methods developed for cell tracking entails two key steps: **(a)** distinguishing the position of the cells from the background and **(b)** linking the individually identified cell in successive image/video frames [2]. These are described as follows:

(a) Conventionally, there are two main image segmentation approaches which are based on *thresholding* [2–4] and *template matching* [5]. At its simplest form, the former entails choosing a threshold, such that any pixels above it are cell and any below are background [6]. More sophisticated thresholding methods rely on the additional use of the so called morphological *Watersheds* analysis to help distinguish between cells in highly confluent areas [7]. In particular, the watersheds method has the advantage of working well for cell images, where the background intensity of the image often varies across the *x-y* plane. While this technique is reasonably adept at finding an approximate position of each cell, however, the computed threshold often overestimates the size of a cell since the typically present halo is also included as an integral part of each cell under PC microscopy; see Fig 1. The latter approach to background removal involves finding the position of the foreground objects based on their expected appearance. This is typically achieved by the construction of a cell template using either a collated set of features that best describes the cell/object(s) of interest, or a manually constructed exemplar using samples selected from the first image of each sequence under study. In either case, the approach relies to some extent on the user input which adds time and human error to the process.

(b) A simple method to link the individual cells in successive time frames is based on the *nearest neighbour* principle that seeks to compute a unique label for the individual cell identified previously to its nearest (physical) neighbour in the current frame; see details described in Section 3. In practice, this has a tendency to mislabel cells, especially when they are densely packed in a culture where cells frequently cross paths or even overlap. Alternatively, the linking of the cells can also be achieved by a template matching method that entails aligning the extracted features of the individual cells detected in consecutive image/time frames across the whole experiment. Here, features alignment often depends on an error-free prior segmentation of the image, as any such errors can cause the cells to be incorrectly matched. Consequently, this technique is again most effective with high-contrast images such as those obtained via fluorescence microscopy. Furthermore, any large changes in morphology of a cell captured between consecutive frames may cause it to be labelled as a new cell in the following frame, thus rendering the associated track broken [8].

The remainder of this paper is organised as follows. Section 2 describes the algorithm developed used for this study, highlighting the novel use of *entropy* as a uniform framework to facilitate segmentation and the subsequent linking of cells in successive time frame. Additionally, the investigative method enabling the analysis of vesicles, a contribution of this study, is also presented. The implementation of the tracking algorithm is described in Section 3, which also covers the experimental setup and testing carried out in this study. The results including most relevantly the vesicle analysis are discussed in Section 4, where comparisons with the manually constructed ground truth are discussed. A summary which includes concluding remarks and future research directions are presented in Section 5.

Fig 1: A close up of a back disk cell imaged using PC microscopy at X20 magnification. Showing the two artefacts associated with PC microscopy: the halo, a bright ring around the cell and the shade off, where the edge of the cell is off a similar intensity to the background. The vesicle, the focus for this paper, in PC microscopy it is seen as a small bright circle.

2. METHOD

A new segmentation method that uses textural context to distinguish cells from the background is presented. A principal contribution is the application of the mathematical framework of *entropy* to facilitate development of the proposed cell tracking method that combines the traditionally two-step approach to segmentation and labelling of cells in successive

image frames as described in Section 1 above. In addition, the required data analytics which enable quantitative analysis of vesicles of the individual cells following tracking is also presented.

2.1. Cell Tracking

The key objective the proposed cell tracking algorithm is to compute the locations of the individual cells in each image/time frame and connect for each cell its successive positions in an image sequence. Using entropy measurements, the individual cells are segmented from the background of an image from which the location and area of each cell could be estimated. Such measurements are subsequently used to compute the texture-based feature vector for each cell, enabling it to be connected or mapped to the same cell in successive images, where the required mapping is obtained by means of a cross correlation procedure which is applied to consecutive frames of the image/time sequence.

2.1.1.   Segmentation

The first step used in this paper is to find the entropy of every pixel in relation to its local region, in the predefined 5x5 kernel, resulting in a map as shown in Fig 2. The entropy (E_l) is calculated as:

$$E_l = -\sum p * log_2(p)$$

where p represents the relative histogram count (i.e. the probability) for each image pixel intensity on the 8-bit (0-255) intensity scale as described by Gonzalez et al [9]. Here, E_l calculates the local entropy of an input pixel $l(x,y)$ using a 5x5 kernel defined over the (5x5) neighbourhood of l.

Fig. 2: (Left) A PC image (20X magnification) of back disk cells cultured in 20% Foetal Calf Serum. (Right) The corresponding local entropy ma shows low values of local entropy in grey and high values of local entropy in white. It is noted that cells are identified as objects in the foreground with relatively high entropy values.
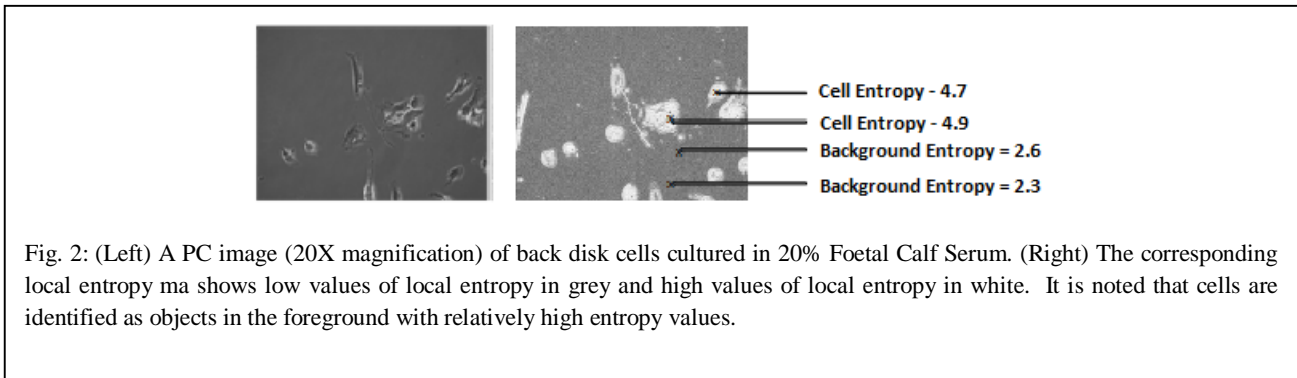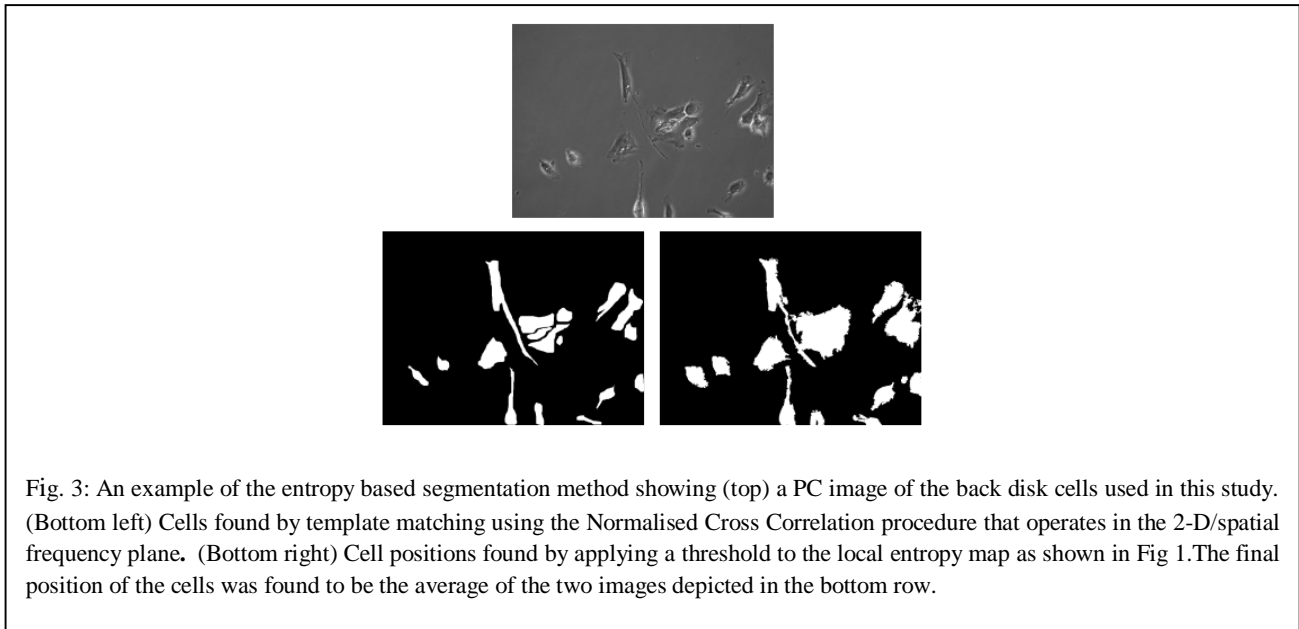
Figure 2 illustrates how the local entropy/map was used to split an image into 2 distinct groups, with low entropy representing the background and high entropy is regarded as foreground cell/objects. A relatively simple thresholding method can then be applied to the entropy map to extract these objects, as pixels with entropy values less than or equal to T are considered as background. More importantly, a key advantage of this method over the Watersheds based segmentation technique is that the halos associated with the individual cells are largely removed as such structures consist of relatively uniform pixels which have low local entropy values. As such, the method is consistent with the basic principle of entropy encoding for which data streams are compressed losslessly as exemplified in the widely known JPEG image format [9]. An example of this segmentation approach is depicted in Fig. 3, where the specialised template matching procedure which links the individually segmented cells over consecutive image/time frames is also described.

2.1.2.   Template matching

The linking stage of the proposed tracking algorithm is divided into two steps. The first is to compute a *feature vector (f)* for each cell which has been identified previously following the entropy based segmentation procedure. This utilises the entropy estimated for each cell and augmented it with additional textural features to include *smoothness* and *contrast*

plus the spatial location of the cell in the form of a feature vector f. A weighting vector w is then applied to f using the standard dot product ($f.w$) that results in a scalar feature score for each cell. As discussed above, this feature score provides a unique descriptor of the individual cell identified in each image and is used to facilitate its mapping of and subsequently linking to the same cell in the next frame. Based on the experiments carried out in this study, the individual scores obtained did not change by more than ±0.2, or 20%, between two consecutive image/time frames (where the scalar dot product is normalised). As such, they offer a reasonably reliable metrics by which an initial list of possible mappings could be obtained for any two consecutive images in the image sequence captured over the course of an experiment.

Fig. 3: An example of the entropy based segmentation method showing (top) a PC image of the back disk cells used in this study. (Bottom left) Cells found by template matching using the Normalised Cross Correlation procedure that operates in the 2-D/spatial frequency plane. (Bottom right) Cell positions found by applying a threshold to the local entropy map as shown in Fig 1.The final position of the cells was found to be the average of the two images depicted in the bottom row.

In passing, it should be noted that this initial list of cell mappings that link the individual cells of the current image to those obtained in the previous frame is either incomplete or non-unique (often) due to the fact that cells (new or otherwise) could enter into or move out of the current field of view. In the case of a new cell moving into the current frame, a new labelled must be added to the list, while any cells which have moved out from the previous frame must be withdrawn (but not removed) from the current list of tracked cells. To address this and other anomalies arising from the ambiguity of the similar feature scores that are within the predefined threshold (±0.2), the initial list of cell assignments must be refined to facilitate a one-to-one mapping of the individual cells obtained between any two consecutive frames. Here, two approaches could be adopted; namely, the *Sum of Absolute Difference* (SAD) [**10**] and *Normalised Cross Correlation* (NCC) procedures [11-13], which are summarised below.

The first approach based on SAD considers the intensities changes across the pixels representing the individual cell in the current frame and find the best match from the preceding frame. It works by constructing a template that covers the 2-D region of the individual cell of interest and finding the best match for this template/region in the preceding image frame. The latter is achieved by computing the minimum of the sum of the absolute pixel by pixel differences of all possible regions covered by the template within the preceding frame. Here, while the SAD procedure is conceptually simple to implement, it could present a major computational bottleneck in practice as the SAD calculation must be performed for each pixel of the image for which the minimum (and hence the best match) is to be found [10]. More importantly, any significant changes in morphology of a cell captured between consecutive frames may also render such a procedure completely inadequate [8].

Alternative, the second approach using the NCC technique seeks to address much of the shortcoming of the template matching based procedure of SAD by using the FFT (Fast Fourier Transform) based convolution technique in the 2-D Fourier domain; see section 3 for implementation notes. The technique has previously been used in a number of image tracking applications [11-1] and, for relatively large target and search images, the FFT based computation is an efficient means to implement correlations, often with several orders of magnitudes faster. Mathematically, the NCC is defined below:

$$NCC = \frac{\sum_{x,y}[f(x,y) - \overline{f_{u,v}}][f(x-u,y-v) - \bar{t}}{\sqrt{\sum_{x,y}[f(x,y) - \overline{f_{u,v}}]^2 \sum_{x,y}[f(x-u,y-v) - \bar{t}]^2}}$$

where $f(x,y)$ is the search (or source) image, $\bar{t}$ is the mean of the intensity of the target/image template and $\overline{f_{u,v}}$ is the mean of $f(x,y)$ in the region covered by the source template. As described in [12], the NCC procedure was shown to be able to generate the best position of the target image t in $f(x,y)$ - the search image (in the preceding frame) from which a cell that has a visual appearance extremely similar to t could be located. This was borne out by the experiments developed for this study.

2.2. Quantitative measurements of vesicles

The final yet important objective of the proposed method concerns the quantitative analysis of vesicles that are present in the individually tracked cells over the course of our experiments. Here, given the vesicles of interest are most noticeable by their relatively bright appearance within each tracked cell (see Fig. 1), a relatively simple thresholding procedure has been applied to facilitate generation of the relevant 2-D masks for the relevant cell in order that quantitative measurements such as size and number of the detected vesicles can be obtained – *viz.*, the *granulometric* studies of vesicles [9,14].

In passing, it should be noted that further studies of vesicles concerning automated analysis of the size and distribution of vesicles is beyond the scope of this paper. Interested reader should also be referred to our allied work in progress that investigates in detail causal links between vesicle formation and the survival mechanisms of intervertebral disc[1].

Finally, as with the cell tracking algorithm presented above, the results obtained for this part of our work has also been validated against the manually constructed ground truth; with vesicle counts generated by hand using a freehand drawing tool. This is further discussed in Section 4 below.

3. IMPLIMENTATION AND TESTING

The microscopic image/data sequences studied here were obtained using the high-throughput time lapsed microscope[2] equipped with a high precision motorised *z*-stage that enables the automated acquisition of the so-called *all-in-focus* images frame-to-frame over time. Coupled with the use of PC optics, the imaging platform allows the integration of high speed and continuous focusing software to keep cells in focus before their acquisition in a fluorescent free environment. These images are then studied and analysed using customised MATLAB® analysis code developed in-house, which segments and tracks cells frame-by-frame over the complete course of the experiments developed for this study. In essence, the tracking implemented connects cells from one frame to the next by matching the texture based feature scores

---

[1] S Brown et al, *Mechanisms that intervertebral disc cells utilise to protect them in their challenging environment*, under review in Cells and Materials Journal (ecmjournal.org).

[2] Cell-IQ2®, CM Technologies Ltd.

computed for the individually segmented cells and between cell positions in consecutive images. Here, it should be noted that, when a new cell is detected in the current field of view (e.g. a cell division event is identified), the tracker would also attempt to connect parent and child cells by minimising the distances between the two child cells and the distance between the child and parent cells. To facilitate efficient tracking of vesicles development, the algorithm followed the individual cells on the basis of their dynamic *centroid* positions. Once completed, this allowed further studying of the identified cells, including cell counting, area measurements and, relevantly, the detection and quantitative analysis of the vesicles in each of the individually tracked cells.

Several comments are in order. First, the template matching procedure described in Section 2.1.2 concerns initially aligning the feature vector (**f**) based scores generated for each segmented cell over consecutive image frames of the time-lapsed video sequences. This is followed by the compute intensive task of resolving any ambiguities arising from the similarly scored cell objects that are present in close proximity commonly occurs in high cell/confluent areas. Such a feature matching approach adopts the widely known image/signal correlation technique that locates in the current frame (or more appropriately, the current or *target* field of view) a specific or *source* cell/pattern identified in the preceding image frame. Computationally, the required correlation operation is very similar to convolution, with the amplitude of each sample/pixel in the cross-correlation image a measure of how much the source *resembles* the target image at that location. In other words, the value of the cross-correlation is maximised when the target image/signal is aligned with the same features in the source image/template. As with most image convolution procedures, a large number of calculations were required involved and this problem was addressed by making use of the well documented technique of FFT convolution[3], upon which the implementation of the proposed NCC procedure was based. Second, the granulometric studies of vesicles were facilitated using a somewhat elaborated implementation of morphological analytics that measure particle sizes distribution indirectly; *i.e*., without identifying explicitly and measuring every particle. For vesicles with generally round shapes that are predominately brighter than the enclosing cell body, the basic approach is to apply morphological openings of increasing size whereby the sum of all pixel values in the opening is computed as the surface area covered by a vesicle for each opening. The ensuing analysis thus included measuring the size and shape of the vesicles detected and their arrangement within each cell. Using the PC microscopy where unlabelled cells were individually identified, the morphological parameters could also be determined for each cell, allowing the same set of operations to be applied on the whole image and/or on specific regions of interest where vesicles were detected and, importantly, enabling statistically robust analyses to be performed on the necessarily large number of cells.

To facilitate testing and validations of the proposed cell tracking algorithm, thirty cells (see below) were selected and compared with the results obtained through manual tracking of these cells over the course of the study. Here, the first ten frames of the microscopic image sequence containing the vesicles were selected manually by an expert and they were subsequently applied to the tracking algorithm in this paper. To determine the accuracy of the segmentation and tracking the parameters listed below were selected:

1. Number of cells correctly located (in each of the ten frames).
2. Number of cells missed by the automated tracking.
3. Number of false positives. That is, the current cell was identified in the preceding frame but was mapped to in a region which contains only background/noise, or with multiple cell centres that has been marked incorrectly as a single cell.
4. The number of cells successfully tracked such that they had been correctly located in each of the 10 frames from start to finish. Here, if a cell was located at a different position to the manually marked position on the last frame of the sequence, or not located at all, the track was considered to be lost. Only the cells that were correctly located on the first frame were included in these results in order that an objective comparison, independent on the initial segmentation, could be obtained**.**

---

[3]See a complete example described at http://uk.mathworks.com/help/images/fourier-transform.html

It should be noted that, in any case, a cell was defined as correctly segmented if it contained 90% of the same pixels as those of the manual segmentation. A cell was said to be tracked accurately if it maintained the same label throughout the ten frames for which it was being tracked. To ensure that the tracking algorithm was working correctly across the entire 10-image sequence, three separate sets such image sequences were analysed. These were drawn, respectively, from (a) the start (0% mark); (b) half of the way through (50% mark); and (c) towards the end (90%) of the selected image sequence.

Finally, similar tests were also carried out in the context of vesicle analysis, including:

1. Number of vesicles correctly detected.
2. Number of vesicles missed; *i.e.* no vesicles were identified within the associated cell.
3. Number of false-positives; *i.e.* vesicles were identified in area that contains only background noise or, occasionally, multiple cell centres had been marked as a single cell.

As with the proposed cell tracking algorithm, the results obtained were compared with the manually marked positions of the vesicles that were present in each of the correctly tracked cell as described above. For simplicity, a single position in the centre of each detected vesicle was taken to be the reference location, for which a vesicle was deemed to be correctly labelled if the enclosing cell contained such vesicle location.

4. RESULTS AND DISCUSSION

A computational workflow was developed to automate tracking and analysing potentially thousands of cells captured using high-throughput time lapsed microscopic imaging platform (CellIQ2). The tasks at hand were complex and entailed the development of the presented tracking algorithm (in Section 2) which sought to locate the cell/features of interest within the culture, by computing, firstly, the position of the individual cell and, secondly, the relative locations of vesicles identified in each of the individually tracked cells over the course of the experiment. Specifically, the techniques adopted help quantify such physical movements within the field of view, by separating from the image background the individual cell in the entropy space where the locally measured entropy was also added to the dynamically constructed texture features descriptor of each cell across the time lapsed image sequence. Coupled with the NCC procedure that was designed to operate in the frequency domain, such a descriptor enables individual cell to be tracked and, importantly, analysed for the presence of vesicles over time.

Several tests have been devised to help develop/validate the proposed tracking algorithm that was accompanied by closely allied data analytics specifically designed to analyse vesicles development over the course of our study. These tests have been implemented using a manually construct ground truth (GT) set which can be defined as follows:

1. For cell segmentation; the GT consists of the manually validated location of each cell selected.
2. For cell linking; the GT consists of the complete mappings of the individual cell over successive time frames.
3. For vesicle quantification; the GT consists of vesicle counts for the individually tracked cell over the selected time frames.

4.1. Validations: cell segmentation and linking

Figure 4 summarises the results obtained by the automated cell segmentation of cells were compared with the GT set (consisting of a total of 87cells), where it was found that 87% of the cells were found/labelled correctly, 5% were found to be background and/or debris which have been incorrectly labelled as cells (*i.e.* false positives) and the remaining 8% were without a label (*i.e.* false negatives). The latter includes mostly those that represent a large group of closely populated cells were labelled as one cell. These results were further compared to that generated by the *Cell Tracker*[4], a

---

[4] *CellTracker*: a program for automated cell tracking on phase contrast images. Available at http://group.szbk.u-szeged.hu/sysbiol/horvath-peter-lab-overview.html

freely available software platform that combines the techniques of particle filtering and active contours technique to determine the relative position of the nucleus and cytoplasm of the individual cell. Using the same test data and GT set, the cell tracking platform was only able to correctly label 50% of the total number of cells, although with fewer false positives (2%) when compared with the proposed algorithm (5%).
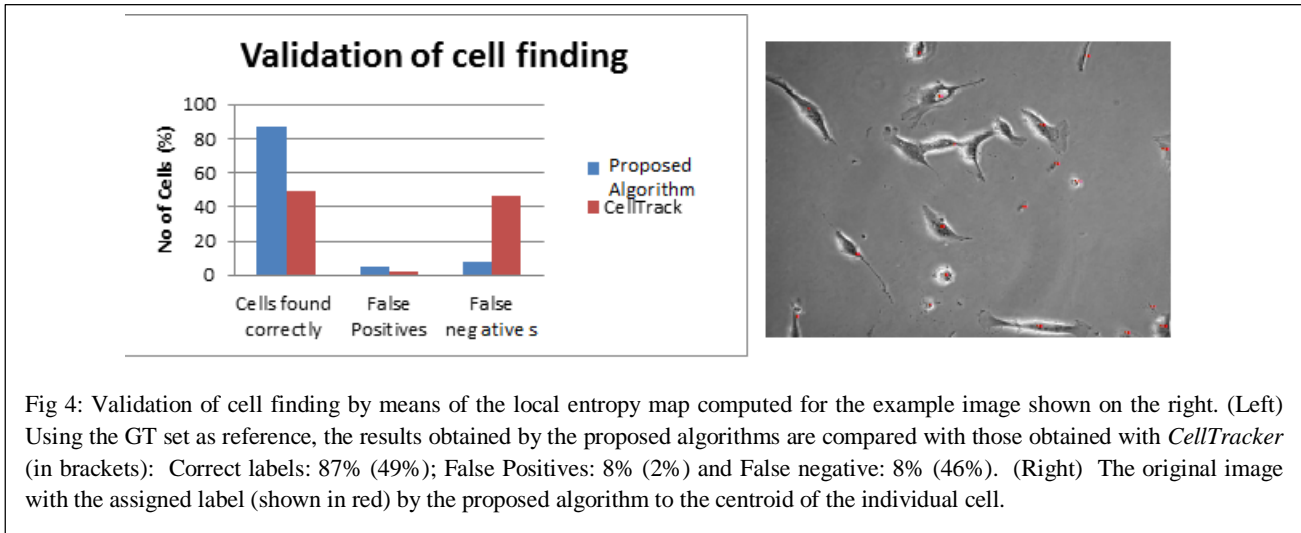

Fig 4: Validation of cell finding by means of the local entropy map computed for the example image shown on the right. (Left) Using the GT set as reference, the results obtained by the proposed algorithms are compared with those obtained with *CellTracker* (in brackets): Correct labels: 87% (49%); False Positives: 8% (2%) and False negative: 8% (46%). (Right) The original image with the assigned label (shown in red) by the proposed algorithm to the centroid of the individual cell.
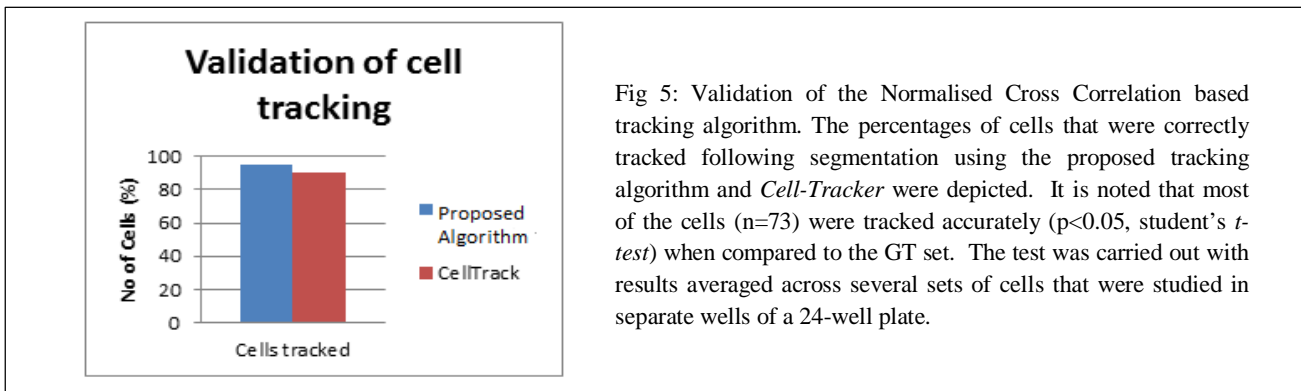

Fig 5: Validation of the Normalised Cross Correlation based tracking algorithm. The percentages of cells that were correctly tracked following segmentation using the proposed tracking algorithm and *Cell-Tracker* were depicted. It is noted that most of the cells (n=73) were tracked accurately ($p<0.05$, student's *t-test*) when compared to the GT set. The test was carried out with results averaged across several sets of cells that were studied in separate wells of a 24-well plate.
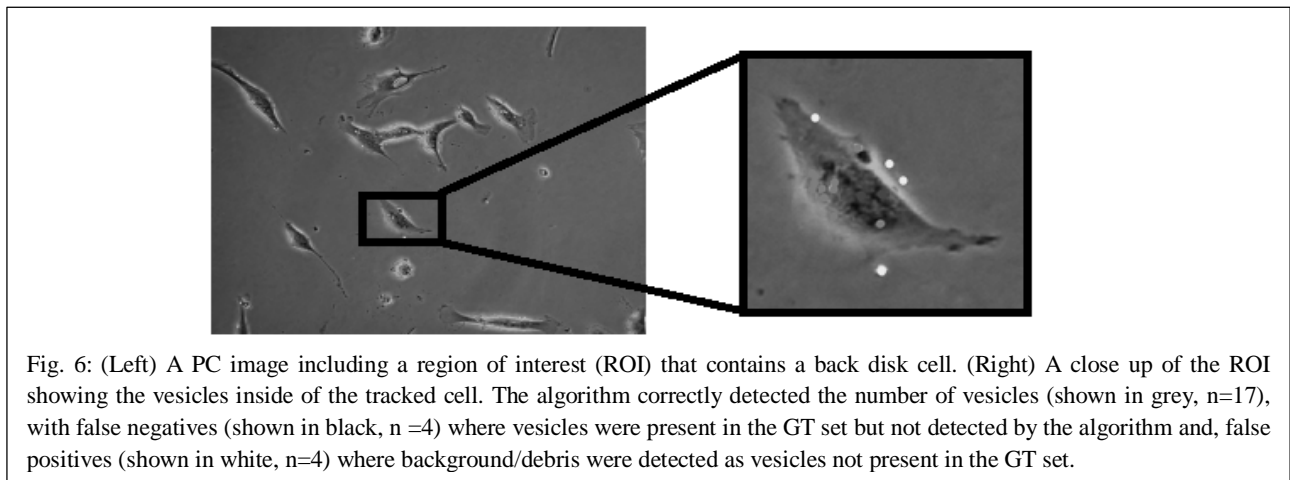
It was also found that the proposed NCC based tracking algorithm was able to track the cells accurately, with $p<0.05$, using a null hypothesis that assumed the individual cells would not be tracked or linked from the first to the last or tenth image of the selected sequence. This test was carried out with results averaged across several sets of cells that were studied in separate wells of a 24-well plate. Only cells that were correctly identified in the first image were used in this test, ensuring that the results obtained by the automated tracking procedure would not be dependent on the results of the initial segmentation. Cells were deemed to have been track correctly if they also retained the same numerical feature score for the first and the last (10th) image when compared to the manually derived score values. It is evident in Fig. 5 that the majority of cells were correctly tracked by the automated tracking algorithm when compared to the manually constructed GT[5]. It is also noted that cells that were incorrectly tracked were often occluded significantly (> 50%) by the other cells. These results were found to be comparable to that obtained by Cell-Tracker, which tracked 89% of the

---

[5] Measurements were taken from a freehand drawings (CROIEditor) overlaid on the all-in-focus image obtained via the high-throughput time lapsed microscopy platform (Cell-IQ2).

individual cell correctly. However, as Cell-Tracker was only able to identify correctly 50% of the cells, initial pool from which to draw these cells for subsequent tracking was much smaller.

4.2. Vesicle Analysis

This last but important study was designed in a specific biological context which investigated vesicle development in back disk cells over time. Specifically, it sought to capture minute dynamic details of these cells, with a 20x lens to facilitate quantitative analysis of vesicles present in the individual cell. Here, a key objective of the developed algorithm was to facilitate generation of biologically meaningful measurements of back disk cells that are based on automated quantitative studies of vesicles developed in these cells over the course of the experiments. As proof of principle, the data analytics were so constructed that (a) they would enable vesicles to be identified reliably within the individually tracked disk cells across successive time frames and, more relevantly, (b) the statistical distribution of back disk cells could be characterised based on the vesicles present in these cells in a simple yet biologically relevant context. For (b), a preliminary test was carried on one of selected image sequences where the individual cells were previously tracked. This is illustrated with results presented in Fig. 6.

Fig. 6: (Left) A PC image including a region of interest (ROI) that contains a back disk cell. (Right) A close up of the ROI showing the vesicles inside of the tracked cell. The algorithm correctly detected the number of vesicles (shown in grey, n=17), with false negatives (shown in black, n =4) where vesicles were present in the GT set but not detected by the algorithm and, false positives (shown in white, n=4) where background/debris were detected as vesicles not present in the GT set.

A further experiment was developed to help track/study the different numbers or levels of vesicle in these cells. This was achieved by splitting the individually tracked cells into three distinctive groups based on the number of vesicles contained within each cell: namely, cells with fewer than three vesicles were assigned to Group-A, 3-10 vesicles to Group-B and over 10 vesicles to Group-C. The results are presented in Fig. 7, where it is apparent that automated algorithm produced vesicle counts which were broadly in line with the manually constructed GT set, with 87% of vesicle correctly identified/labelled, 5% false positives (caused by background and debris) and 8% false negatives (attributed to groups of smaller vesicles which were labelled as a single vesicle).

More specifically, when compared to the GT set where the vesicles contained in the individual cells were manually counted, the results generated by the automated algorithm also varied across the three vesicle groups; viz., for Group-A, the algorithm (solid line) produced a noticeably higher estimated cell counts than the GT set (dashed line), whereas the estimated cell counts are only slightly higher for Group-B and are indeed consistently lower for Group-C. In the latter two cases (Group-B and Group-C), however, the number of detected cells were far too small (<10) to enable any statistically significant or meaningful conclusions to be drawn. This warrants further investigations.
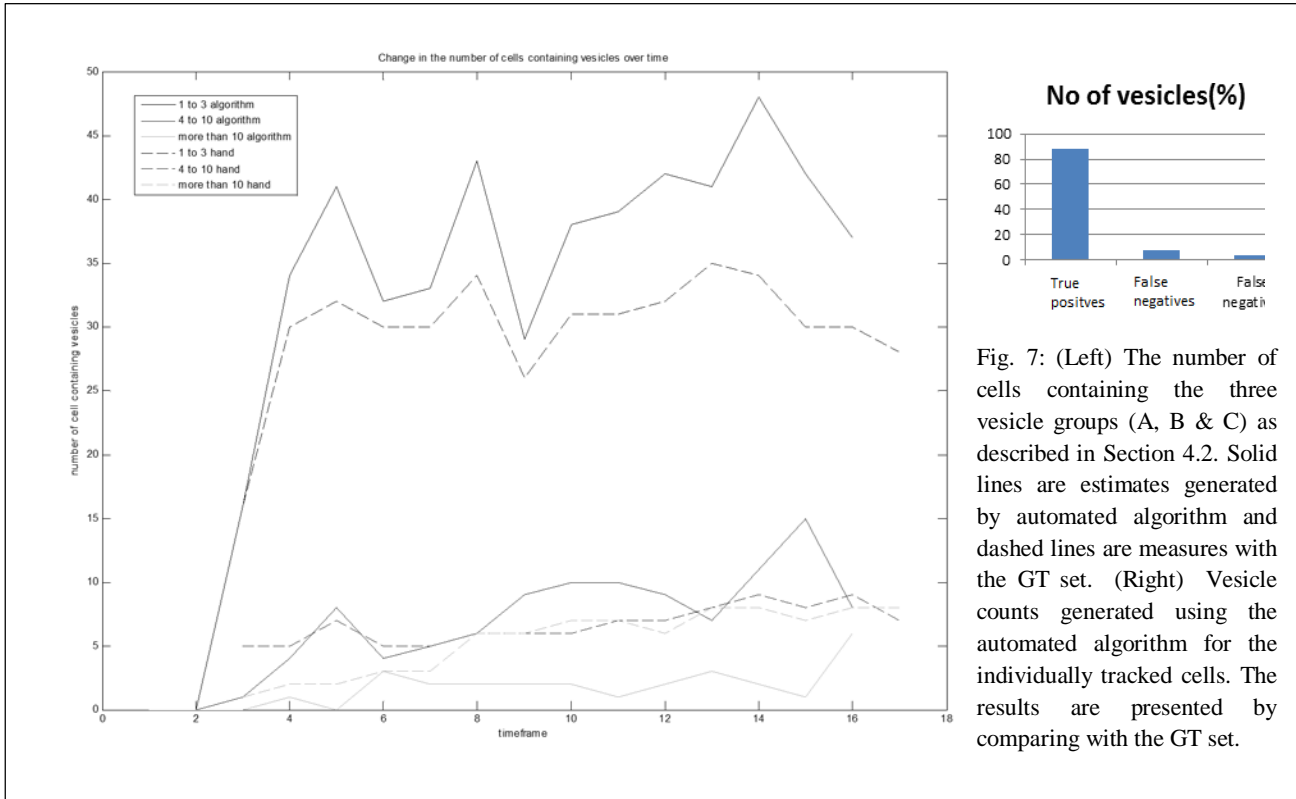
Fig. 7: (Left) The number of cells containing the three vesicle groups (A, B & C) as described in Section 4.2. Solid lines are estimates generated by automated algorithm and dashed lines are measures with the GT set. (Right) Vesicle counts generated using the automated algorithm for the individually tracked cells. The results are presented by comparing with the GT set.

The final test was designed to investigate the impact or influence of cell sizes on vesicles development. Here, the cells were split into two groups; those with more than three vesicles and those without. As discussed above, cells with less than three vesicles (Group-A) were discarded to minimise the likelihood of false positives as generated by the proposed tracking algorithm. The results are summarised in Fig. 8, where it is shown that, contrary to in our intuition, vesicles were mostly found in the smaller cells, with p<0.001 when the Student's *t*-test was applied using the mean value of cell sizes which were measured as the total number of pixels covered by the individual cells. This finding is of particular interest; until now, the relatively simple study of the effect of cell sizes on vesicles development has not been possible without performing the often laborious and (thus) inevitably error-prone manual measurements.
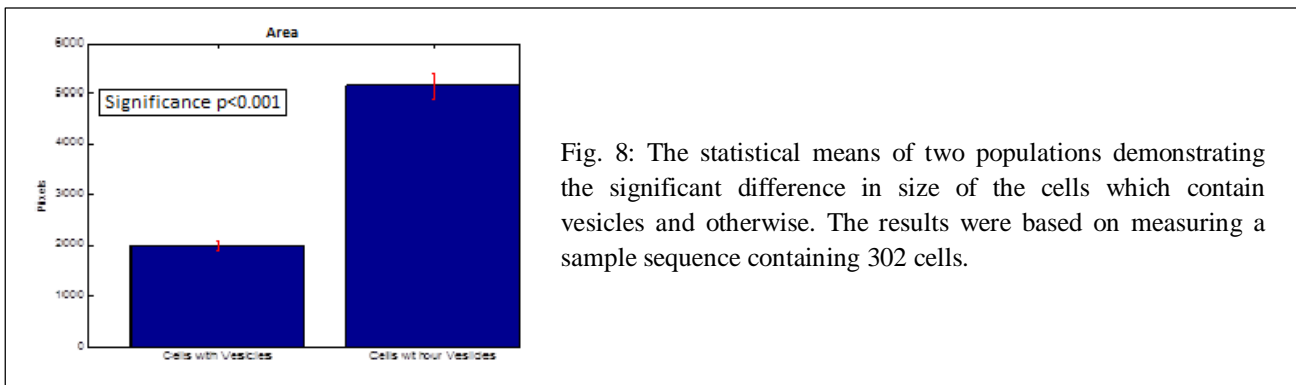

Fig. 8: The statistical means of two populations demonstrating the significant difference in size of the cells which contain vesicles and otherwise. The results were based on measuring a sample sequence containing 302 cells.

5. CONCLUSIONS

A new cell tracking algorithm that combines the traditionally two-step approach to segment and link cells is presented. Essentially, the algorithm applies the mathematical framework of entropy to enable the segmentation of and dynamic feature vector construction for unlabelled live cell images captured using high-throughput time-lapsed video microscopy. As proof of principle, additionally, a quantitative analysis of vesicles which are present in the individually tracked cells over the course of our study has also been described. These vesicles were studied in some details using a computational workflow developed to facilitate relevant experimental studies, which sought to test/validate the proposed tracking algorithm and, importantly, the closely allied data analytics which help quantify vesicles development throughout the experiment. Several tests have been implemented to enable biologically relevant results to be studied and compared against the manually constructed ground truth sets. The results show that the procedures developed could reliably track cells in a reasonable time frame, with high success rates and little human/user input. From a computational perspective, they also demonstrate that the combined use of textual features descriptors and Normalised Cross Correlation offers an effective approach to compute the location trajectory of the individual cells throughout the experiment. As such, this enables for the first ever automated studies of vesicles development in back disc cells without using biomarkers and with promising outcomes; viz., *87%* true positives, 5% false positives and 8% false negatives.

As on-going work, the development of advanced data analytics capable of tracking unlabelled vesicles in conjunction with the proposed cell/lineage tracker is currently underway. This would support further collaborative studies on the cellular mechanism of *autophagy* which takes place in the intervertebral disc, the largest avascular tissue in the human body, to improve our understanding of the survival mechanisms of these cells and, perhaps importantly, identify therapeutic targets that may help alleviate degenerative spinal pathologies.

REFERENCES

[1]   A. E. Walsby, "Gas vesicles.," *Microbiol. Rev.*, vol. 58, no. 1, pp. 94–144, Mar. 1994.

[2]   N. Jaccard, L. D. Griffin, A. Keser, R. J. Macown, A. Super, F. S. Veraitch, and N. Szita, "Automated method for the rapid and precise estimation of adherent cell culture characteristics from phase contrast microscopy images.," *Biotechnol. Bioeng.*, Sep. 2013.

[3]   D. G. T. Whitehurst, S. Bryan, E. M. Hay, E. Thomas, J. Young, and N. E. Foster, "Cost-effectiveness of acupuncture care as an adjunct to exercise-based physical therapy for osteoarthritis of the knee.," *Phys. Ther.*, vol. 91, no. 5, pp. 630–41, May 2011.

[4]   Z. Yin, K. Li, T. Kanade, and M. Chen, "Understanding the optics to aid microscopy image segmentation.," *Med. Image Comput. Comput. Assist. Interv.*, vol. 13, no. Pt 1, pp. 209–17, Jan. 2010.

[5]   I. Adanja, O. Debeir, V. Mégalizzi, R. Kiss, N. Warzée, and C. Decaestecker, "Automated tracking of unmarked cells migrating in three-dimensional matrices applied to anti-cancer drug screening.," *Exp. Cell Res.*, vol. 316, no. 2, pp. 181–93, Jan. 2010.

[6]   S. U. Lee, S. Yoon Chung, and R. H. Park, "A comparative performance study of several global thresholding techniques for segmentation," *Comput. Vision, Graph. Image Process.*, vol. 52, no. 2, pp. 171–190, Nov. 1990.

[7]   F. Meyer, "The watershed concept and its use in segmentation : a brief history," Feb. 2012.

[8]   A. J. Hand, T. Sun, D. C. Barber, D. R. Hose, and S. MacNeil, "Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration," *J. Microsc.*, vol. 234, no. 1, pp. 62–79, 2009.

[9]   R. Gonzalez, *Digital Image Processing*, 2nd ed. Prentice-Hall, 2002.

[10]   I. E. Richardson, *H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia*. John Wiley & Sons, 2003, p. 281.

[11]   D.-M. Tsai and C.-T. Lin, "Fast normalized cross correlation for defect detection," *Pattern Recognit. Lett.*, vol. 24, no. 15, pp. 2625–2631, Nov. 2003.

[12]   Feng Zhao, Qingming Huang, and Wen Gao, "Image Matching by Normalized Cross-Correlation," in *2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*, 2006, vol. 2, pp. II–729–II–732.

[13]   K. Norman and R. O'Reilly, "Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach.," *Psycholgical Rev.*, pp. 3–44, 2002.

[14]   B. György, T. G. Szabó, M. Pásztói, Z. Pál, P. Misják, B. Aradi, V. László, E. Pállinger, E. Pap, A. Kittel, G. Nagy, A. Falus, and E. I. Buzás, "Membrane vesicles, current state-of-the-art: emerging role of extracellular vesicles.," *Cell. Mol. Life Sci.*, vol. 68, no. 16, pp. 2667–88, Aug. 2011.