

A Self-Supervised Machine Learning Approach for Objective Live Cell Segmentation and Analysis

Michael C. Robitaille¹, Jeff M. Byers¹, Joseph A. Christodoulides¹, Marc P. Raphael*¹

¹ Materials Science and Technology Division, U.S. Naval Research Laboratory, Washington D.C.

* Corresponding author: Marc.Raphael@nrl.navy.mil

Abstract

Machine learning algorithms hold the promise of greatly improving live cell image analysis by way of (1) analyzing far more imagery than can be achieved by more traditional manual approaches and (2) by eliminating the subjective nature of researchers and diagnosticians selecting the cells or cell features to be included in the analyzed data set. Currently, however, even the most sophisticated model based or machine learning algorithms require user supervision, meaning the subjectivity problem is not removed but rather incorporated into the algorithm's initial training steps and then repeatedly applied to the imagery. To address this roadblock, we have developed a self-supervised machine learning algorithm that recursively trains itself directly from the live cell imagery data, thus providing objective segmentation and quantification. The approach incorporates an optical flow algorithm component to self-label cell and background pixels for training, followed by the extraction of additional feature vectors for the automated generation of a cell/background classification model. Because it is self-trained, the software has no user-adjustable parameters and does not require curated training imagery. The algorithm was applied to automatically segment cells from their background for a variety of cell types and five commonly used imaging modalities - fluorescence, phase contrast, differential interference contrast (DIC), transmitted light and interference reflection microscopy (IRM). The approach is broadly applicable in that it enables completely automated cell segmentation for long-term live cell phenotyping applications, regardless of the input imagery's optical modality, magnification or cell type.

Key Words: live cell imaging, segmentation, phenotyping, machine learning, unsupervised, classification

Introduction

Live cell phenotyping is an information rich experimental approach, capable of providing mechanistic insights into cell biology^{1,2}, guiding drug development³ and elucidating disease pathologies^{4,5}. The wealth of information available from live cell microscopy results from the fact that there are numerous optical modalities that can be integrated within a given experiment – from fluorescence imaging which provides spatio-temporal information on specific signaling pathways and organelles to label-free techniques such as phase contrast and differential interference contrast (DIC) which enable the visualization of whole cellular morphologies and dynamics. Each of these modalities provides its own outcome measures which can be viewed as static snapshots or dynamic variations within the four-dimensional space of x, y, z and time⁶.

However, compared to genotyping - its synergistic partner technique - live cell phenotyping remains a far more subjective science. The generation of genomic sequencing data and its analysis can now be achieved autonomously by employing a combination of robotics and microfluidics for sample preparation and

machine learning algorithms for data collection and interpretation. In contrast, the extraction of quantitative information from live cell imagery by manual means is still commonplace in live cell microscopy, a fact which speaks to the human visual system's adeptness at detecting small changes and low contrast features with high fidelity. But with automated live cell microscopes now able to collect high resolution imagery for days on end, the resulting data files can quickly grow to tens of gigabytes, leaving the analyst with an overwhelming amount of imagery to work through.⁷ Furthermore, if the analyst is not blinded to the experimental design, unconscious bias can creep into the data extraction process.

Enter computational algorithms capable of extracting the relevant outcome variables from the imagery in an automated fashion.⁸⁻¹⁰ Broadly speaking, the algorithms are often classified as model based approaches (e.g. Cell Profiler)¹¹, and machine learning algorithms, (e.g. U-Net, ilastik)¹²⁻¹⁴. Neither approach is completely autonomous when it comes to cell segmentation: model-based approaches require the manual tuning of multiple parameters, while machine learning requires the user provide curated data from which the algorithm is trained. Once tuned or trained, the software is able to process far more imagery than could be achieved manually - but there is still a human-in-the-loop. It is just that the manual contribution has been moved to the front end for training purposes and is then continuously reapplied by the algorithm. Algorithms that are tuned or trained at the onset can problematically miss relevant features as the cellular phenotypes or background characteristics evolve, inadvertently skewing the analysis. For instance, variations in label intensity (e.g. photobleaching, quenching) or new morphological features that were not present during the initial training (e.g. differentiation, mitosis, blebbing) can go undetected if not retrained with a freshly curated data set or parameters that capture the offending features. In the same way, temporal variations in the background illumination intensity or homogeneity can also result in improper cell segmentation.⁷ Especially concerning is that the user-supervised training process is inherently subjective in nature and can cause unconscious biases to be effectively baked in to the extracted data by the training process. To optimize objectivity and efficiency, an essential goal is to develop software that can accept imagery from any optical modality, labeled or unlabeled, and extract the cellular features of interest *without input from the user*.

As participants in a synthetic biology real-time reproducibility project administered by U.S. Defense Advanced Research Projects Agency (DARPA), referred to as Independent Verification & Validation (IV&V), we have recently experienced all of these algorithmic limitations and how they can result in large amounts of data either being incorrectly segmented, subjectively segmented, or left unanalyzed due to time constraints.¹⁵ The program involves a wide range of cell types (amoeboid to eukaryotic) from multiple cell biology laboratories; multiple imaging modalities – both fluorescent and label free; and objective magnifications ranging from 10X to 100X. The cumbersome process of retraining supervised machine learning software to match this variety of conditions proved impractical and a human-in-the loop training step was deemed too subjective. The challenge then was to develop a completely automated segmentation algorithm for live cell microscopy applications. In particular, the image analysis software should be 'self-supervised', meaning it trains itself to classify cells versus background and then regularly updates this training so that it can adapt to evolving intensities and morphologies. The software was required to segment a variety of cell types from live cell imagery given the most common imaging modalities as inputs - phase contrast, transmitted light, DIC, fluorescence and interference reflection microscopy (IRM) – and to do so without user-adjustable parameters or user-selected training imagery. It was additionally required that the generated models adapt to changing cell phenotypes and lighting conditions for long-term imaging applications (hours to days).

Methods

To replace more manual model based and machine learning training approaches for segmenting cells with an automated, self-supervised algorithm, we took advantage of the one phenotypic feature which is present in live cell microscopy no matter what the modality: *motion*. From the nanoscale diffusion of proteins and vesicles to the migration of cells that are tens of microns in length, the ever present dynamics captured by live cell microscopy make it ideal for applying optical flow (OF) algorithms designed to identify not just spatial intensity features in a given frame but also the variation or ‘flow’ of those features from frame to frame. The central assumption in optical flow algorithms is that the overall image intensity will remain constant if the time difference between frames is reasonably small.¹⁶ This leads to the following time-derivative constraint equation:

$$\frac{d}{dt} I(x, y, t) = 0 \rightarrow \frac{dx}{dt} \frac{\partial I}{\partial x} + \frac{dy}{dt} \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0$$

u v

where $I(x, y, t)$ is the in-plane image intensity at time t , u and v being the optical flow in the x and y directions, respectively. The methods used to solve this constraint equation are matched with the imaging goal, such as reducing jitter in imagery taken from helicopters, aligning medical imagery or, in the case of this study, cell motion segmentation. In testing a range of optical flow algorithms for cell segmentation, we found the Farneback method to be the most robust due to its sensitivity to object deformation – a natural fit for cells which are morphologically variable.^{17,18}

OF assumptions may or may not be met for fluorescence time-lapse imagery applications in which extended time intervals are sometimes employed to avoid phototoxicity or photobleaching.^{6,19} For this reason, it was important that our technique be co-validated with label free techniques such as transmitted light and phase contrast which are minimally invasive. Overlays of less frequently accumulated fluorescence imagery with cells segmented using a label-free imaging channel is then straightforward. Furthermore, there has been an increased appreciation for the morphological information label-free approaches can provide as a result of algorithmic-based phenotyping.²⁰⁻²²

Our approach to self-supervised learning and automated model generation begins with utilizing the Farneback OF method as a means of classification bootstrapping (Fig 1). Typical segmentation strategies involve utilizing static information in a single image at time frame (t), which can have difficulty distinguishing ‘cell’ from ‘background’ pixels in a generalizable manner (Fig 1a). In contrast, our approach begins with an OF calculation based on images from consecutive time frames ($t-1$, t). This enables us to leverage the ubiquitous nature of intracellular motion and build a dynamics-based feature vector: pixels with the highest flow are automatically labeled as ‘cell’ pixels, those with the lowest flow are automatically labeled as ‘background’ pixels, and those that do not fit either category remain unlabeled (Fig 1b,c). We note that this automatic self-labeling is broadly applicable in that it is not dependent on principles of any specific optical modality, cell type, or phenotype.

The OF-based self-labeling approach outputs a set of ‘cell’ and ‘background’ labeled pixels which are then used to generate additional entropy and gradient feature vectors at each time point. These static feature vectors are used to train and generate a classifier model which, in the final step, is applied to all pixels in the image for cell segmentation. The algorithm is written in stand-alone MATLAB script and utilizes functions from the Image Processing, Statistics and Machine Learning, and Computer Vision Toolboxes.

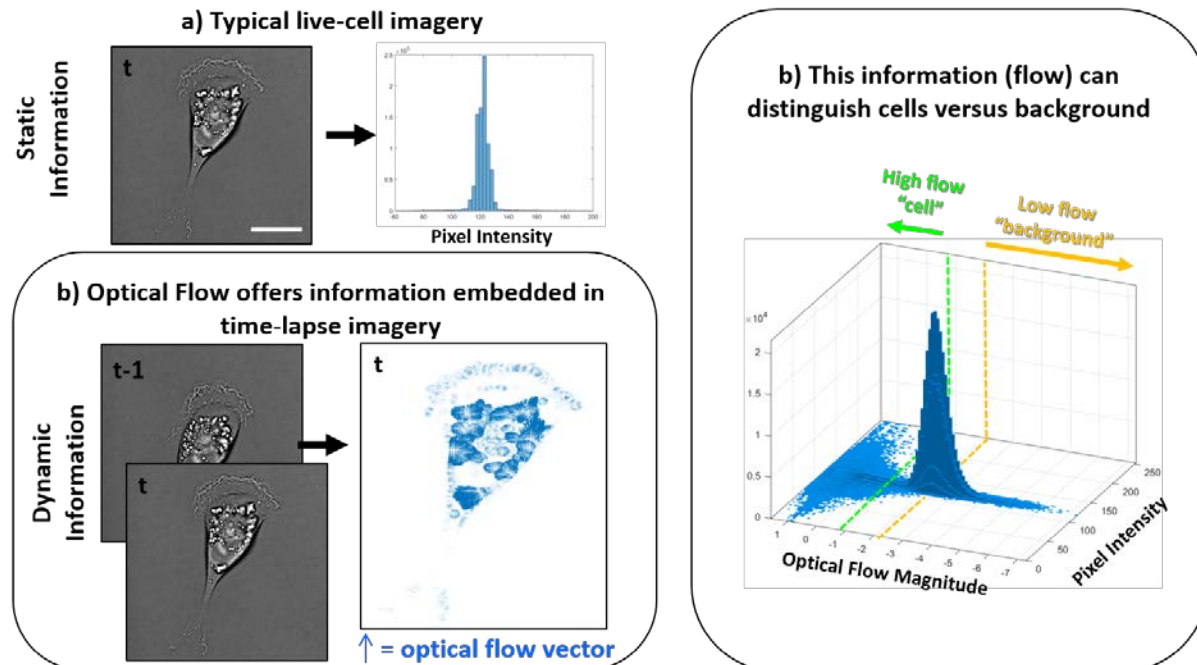


Fig. 1 Overview of the optical flow self-labeling strategy. **(a)** The vast majority of cell segmentation techniques utilize single image frames and the static information contained within as means to distinguish ‘cell’ from ‘background’, oftentimes represented in a histogram. The self-supervised algorithm utilizes optical flow as a means to self-label pixels in an automated fashion. **(b)** Due to the prevalence of intracellular dynamics in time-lapse live cell imagery, optical flow can be calculated for each pair of consecutive images ($t - 1, t$). The optical flow can then be represented as vectors associated with each pixel (right). **(c)** The magnitude of the optical flow then offers a means to distinguish cells from their background, as shown in the bivariate histogram which co-plots the pixel intensity of a single image at t to the optical flow vector magnitudes calculated between consecutive images ($t - 1, t$). Pixels with the highest flow can be automatically labeled ‘cell’ (left of the green dashed line) and those with the lowest flow can be labeled ‘background’ (right of the yellow dashed line). Pixels that do not meet either criteria remain unlabeled, while the self-labeled pixels are used to create a training data set for classification. Time increment: 600 sec, scale bar = 20 μm .

The self-supervised training approach is illustrated in Fig 2 using time lapse DIC imagery of multiple (top) and a single highlighted (bottom) MDA-MB-231 cell. From the raw imagery (Fig 2a,b), many portions of individual cells appear to blend in with the background. However, when the OF self-labeling strategy is applied, the algorithm automatically identifies pixels with high flow magnitude, highlighted as green pixels (Fig 2c,d), which are selected as having the highest probability of correctly being labeled ‘cell’. To automatically label the background, the algorithm over segments, that is, a liberal (low) OF threshold is employed which captures motion from not only the cell but also from nearby background pixels as well. The algorithm sets these pixel values to zero and labels the pixels in which no significant motion was detected as ‘background’ (Fig 2c,d yellow pixels). Once labeled ‘cell’ or ‘background’ in this unsupervised manner by OF (dynamic features from image pair ($t - 1, t$)), entropy and gradient feature vectors (static features from image at t) are generated for each of these training pixels using their local neighborhood of pixels (S.I., Fig S2). These additional feature vectors are then used to train and generate a Naïve Bayesian classifier model which is applied to the entire image in a pixel-wise fashion. The information gained from

the entropy and gradient feature vectors enables pixels which were left unlabeled in the OF training steps (Fig 2c,d grey pixels) to be classified. The contrast enhanced image (Fig 2b) and model-generated segmentation (Fig 2f, teal pixels) show that the algorithm is able to segment the cell with high fidelity (DIC image/segmented boundary overlay, Fig 2g). Importantly, this labeling, training and classifying procedure occurs recursively on each successive pair of $(t - 1, t)$ images, enabling the classifier model to adapt to changing backgrounds and phenotypes. By using optical flow to label the highest flow pixels as 'cells' and lowest flow pixels as 'background', the labeling process has become automated (or 'self-supervised') and no manual inputs or training images are needed.

For extremely low contrast imagery there can be too few training pixels labeled 'cell' for robust segmentation to occur given the initial OF threshold setting. In such cases, the algorithm calculates the entropy associated with 'cell' pixels and iteratively reduces the OF threshold until the associated 'cell' entropy feature vector is well distinguished from that of the 'background' entropy feature vector.

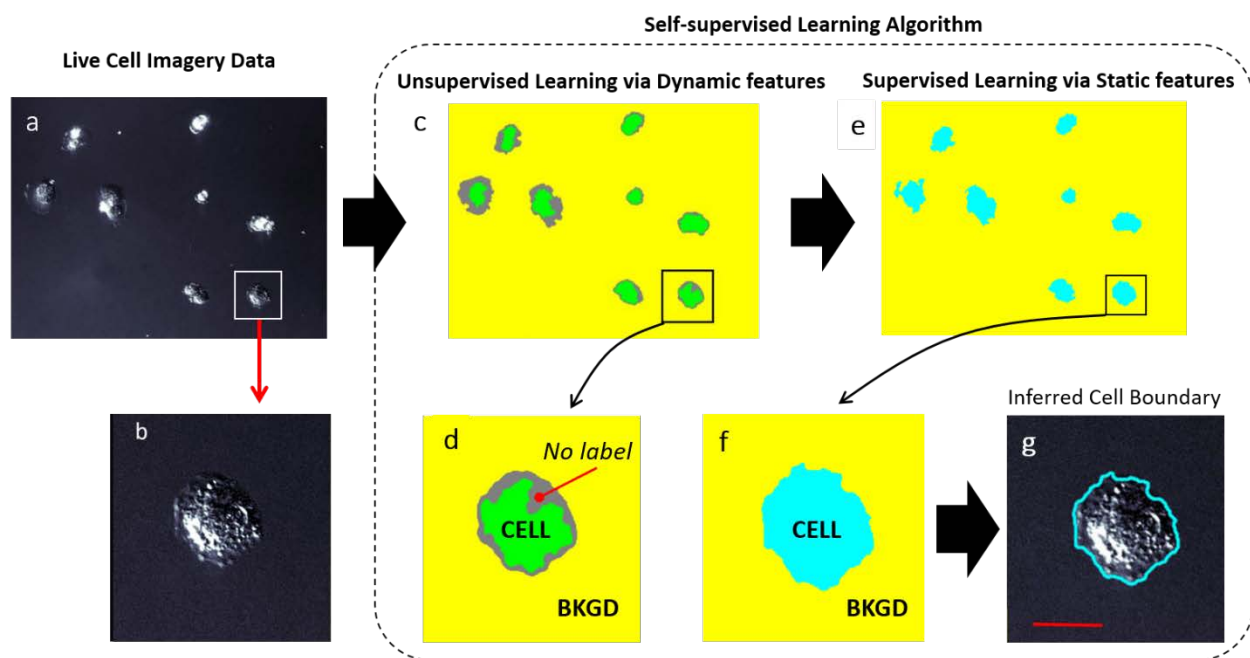


Fig. 2 Overview of the automated self-supervised learning algorithm. **a.** The contrast enhanced DIC image of several and **b** a single highlighted MDA-MB-231 cell illustrates the range of intensities inherent within the cells. (20X objective). **c. & d.** Unsupervised learning via OF: high threshold OF is used to select only those pixels exhibiting the highest flow magnitudes and labels them as 'cell' (green pixels). Similarly, low threshold OF is used to identify pixels with a much wider range of flow magnitudes than the high flow regime. The lowest flow magnitude pixels are labelled 'background' (yellow pixels). Pixels that exhibit OF in between these regimes remain unlabeled (gray pixels). **e. & f.** Supervised learning via self-labeled training data. The self-labeled pixels (green and yellow) are then used to generate static feature vectors, which are in turn used to train the classifier model. **g.** The blue outline is the resulting segmentation which outlines all pixels classified by the OF trained model as 'cell' and is also overlaid on the image in **b**. This process is repeated at every time step, thereby using the most recent imagery to update the training data. Scale bar: 25 μ m (20X objective, time increment: 300 sec).

Results

The Fig 3 imagery shows the generality of this approach and also demonstrates how the self-supervised algorithm additionally automates commonly required manual inputs such as size filtering and hole filling. The segmented cells were processed from imagery acquired from a range of cell types, imaging modalities, magnifications and time increments (S.I. Table S1). The OF algorithm enabled a straightforward approach to automated size filtering which is a common user adjustable parameter in supervised machine learning approaches. To accomplish this, a stand-alone application of OF was applied to the imagery which lacked the added steps of self-tuning and model building described above. While some cell features are missed, this simpler, faster approach was found to be more than precise enough to estimate average cell size and to exclude much smaller objects, thus automating the size filtering process. Because extraneous debris often lacked the motion of the live cells, this debris was also automatically labeled as background by the OF algorithm. Fig 3a and b demonstrate the self-supervised code's ability to size filter, while also adapting to cell types of differing sizes, by comparing the segmentation of human fibroblasts (10X, phase contrast) to those of the much smaller Dictyostelium amoeboid cells (10X, transmitted light), respectively. Extraneous debris features in the Hs27 imagery (Fig 3a, white arrows) are correctly identified as 'background', even though similar in size and intensity to the Dictyostelium cells of Fig 3b. The background inhomogeneities observed in Fig 3a and 3b, which could potentially be mislabeled as 'cell', are correctly identified because they remain relatively constant from frame $t - 1$ to frame t . The segmentation results of the MDA-MB-231 cells (10X, phase contrast) in Fig 3c illustrates the algorithm's ability to adapt to a wide range of phenotypes, from rounded Fig 3c(i) to spread Fig 3c(ii), which is enabled without need for user input by continuously retraining the model on consecutive image pairs. The current instantiation of the software does not attempt to separate cells that are touching or close enough to be segmented as a single object. Well-developed approaches such as watershed transforms²³ and levelset methods²⁴ can be employed for such purposes.

The algorithm works robustly for a range of optical modalities and magnifications as shown in Figs 3d-f. Figs 3d and 3e are segmentation results from IRM imagery (40X, Hs27 cell) and DIC imagery (20X, MDA-MB-231). As a fluorescence imaging example, a self-supervised segmentation of a GFP-actin labeled A549 cell at 100X magnification is shown in Fig 3f. As an additional option, OF can be applied not only as an algorithm labeling element, but also a measurement tool, as shown in the Fig 3f vector plot. The plotted OF vectors (blue) display the magnitude and direction of the measured GFP labelled actin flow between frames. Such measurements have been shown to be useful for quantifying intracellular protein and calcium signaling dynamics.²⁵⁻²⁷

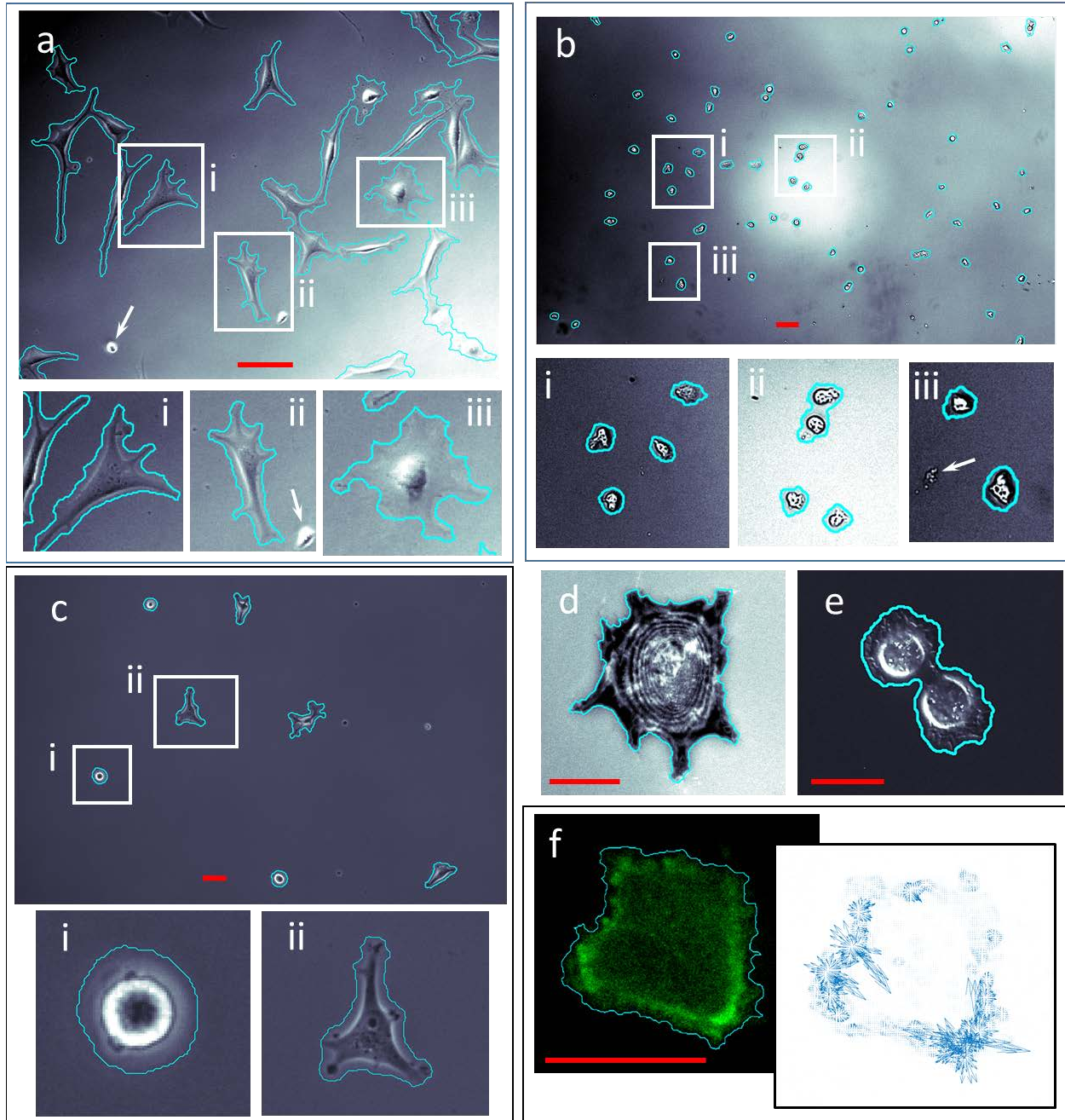


Fig. 3 Self-supervised segmentation for a range of cell types, microscope modalities, time resolutions and magnifications. **a.** phase contrast of Hs27 fibroblasts (10X objective, time increment: 1200 sec) **b.** transmitted light of Dictyostelium (10X objective, time increment: 60 sec) **c.** phase contrast of MDA-MB-231 (10X objective, time increment: 600 sec) **d.** IRM image of a single Hs27 cell (40X objective, time increment: 600 sec) **e.** DIC image of MDA-MB-231 cells (20X objective, time increment: 120 sec) **f.** fluorescence image of a single lifeAct (GFP-actin conjugate) transfected A549 cell (pseudocolored) with the associated optical flow vector plot (100X objective, time increment: 10 sec). Insets i, ii, iii highlight boxed image regions. White arrows point to examples of debris that was correctly labelled 'background' due either to lack of motion or automated size filtering. Images have been contrast enhanced to highlight low contrast features and background inhomogeneities. DIC image (e) was additionally enhanced with a

sharpen filter to highlight interference induced shadowing of cell features. Scale bars: a, b, c: 50 μm ; d, e: 25 μm ; f: 10 μm .

Hole filling, another often required manual input for model-based and machine learning algorithms, has also been automated by this approach. Common examples of when hole filling input is required include fluorescent labels that do not penetrate the nucleus or, for label-free microscopy modes such as phase contrast, large spread cells in which the algorithm has a difficult time associating the interference enhanced cell edges with the enclosed lamellipodia. We found that motion within cells was ubiquitously detected by OF, regardless of imaging modality or whether imaging the cell membrane, nucleus or cytoplasm. Because motion detection was far more common than not for a given pixel within an area labeled 'cell', a fixed morphological blurring tool (circular with a radius of 5 pixels) was found to robustly hole fill regardless of cell type or microscope configuration. The calculated cell area was found to be invariant for a range of blurring tool radii (Fig S2). In all cases, the use of optical flow to identify motion and the 5 pixel radius blurring tool was sufficient to correctly fill in the cell.

By re-training on every pair of consecutive images the self-supervised algorithm remains accurate throughout long-term imaging applications, despite changes in background or cell phenotypes. This allows for a rich behavior of dynamic morphology and migration to readily be collected and analyzed – a key point given the known inter-relationship between cellular shape and function.^{2,28,29} Furthermore, the emerging role that not just cell shape, but cell shape dynamics play in fundamental biological processes is becoming increasingly clear.³⁰

Fig 4 demonstrates how such quantitative morphological information is readily mined in a long-term imaging application. Fig 4a-c shows the tracking of several MDA-MB-231 cells segmented via the self-supervised approach under 10X phase contrast microscopy on cRGD functionalized gold coverslips.³¹ Fig 4a shows the labeled tracks of the cells' centroids over the course of 400 minutes, with the corresponding initial and final image shown in Figs 4b,c. The cell associated with track 2 undergoes mitosis at approximately 320 minutes, creating two new tracks (5 and 6) for the daughter cells. Because the self-supervised approach automatically re-trains continuously on consecutive frame pairs, the morphological changes from Fig 4b to Fig 4c are quantified with high fidelity, as can be seen by plotting the segmented boundaries as a function of time (Fig 4d).

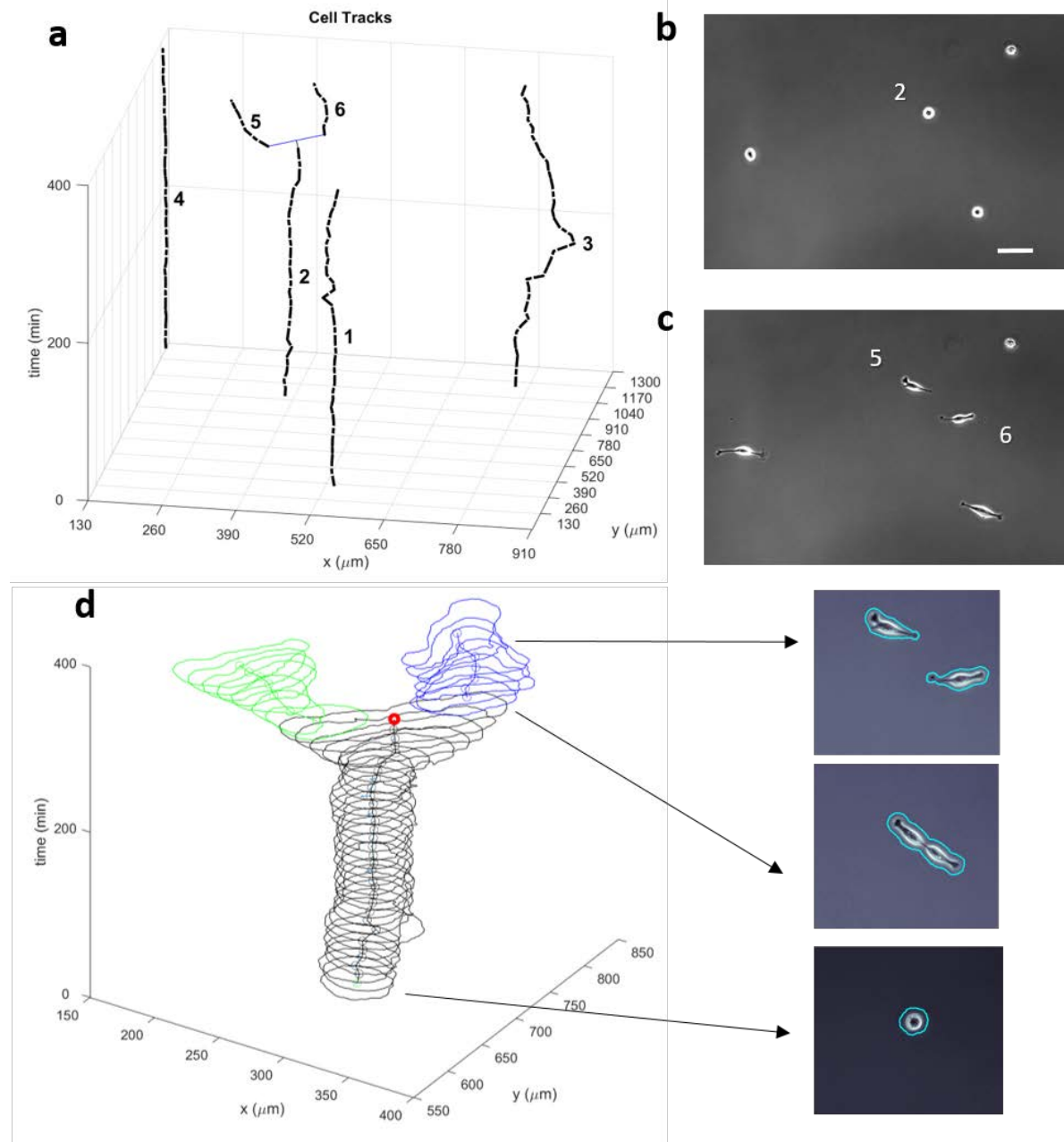


Fig. 4 Tracking of MDA-MB-231 cells under 10X phase contrast microscopy and time evolution of cell morphology through mitosis. **a.** The resulting tracks of multiple segmented cells from a single field of view over the course of 400 minutes **b.** corresponding images at times $t = 0$ min and **c.** 400 min. Track 2 undergoes mitosis resulting in tracks 5 and 6 of the daughter cells (blue line). **d. (left)** Time evolution of segmented morphology of track 2 (black) with the centroid of each shape denoted by an open circle until mitosis, after which the track splits into 5 (green) and 6 (blue), with the cell separation event denoted by a single red open circle. **d. (right)** selected images showing raw data overlapped with the self-supervised segmentation throughout mitosis event. (10X objective, time increment: 600 sec) Scale bar: 100 μm .

Discussion & Conclusions

There are numerous advantages to this self-supervised machine learning approach. The most obvious is that because the training data is generated by tracking motion, the approach can be used with any live cell imaging microscopy technique, whether labeled or label-free. Also unique is the use of the optical flow labeled pixels to self-supervise the building of a classifier model, which in turn is modular with regards to the incorporated feature vectors. While we have employed only two feature vectors in this current instantiation of the classification code (gradient and entropy), there are many additional image features that can be added based on the application. We have also shown that the incorporation of OF enables the straightforward automation of morphological operations such as size filtering and hole filling, eliminating the need for manually tuning these parameters.

The automation described here is markedly different from machine learning approaches that require user assisted training. The most time consuming aspect of model-based tuning and machine learning approaches is the training process. The process is one of trial and error, requiring retraining if the model's performance is not deemed adequate. The complete automation of both the training and segmentation algorithms not only saves time but also removes the chances of unconscious bias from entering the training process. Because the training is conducted recursively with each new image, evolutions in phenotype and background structure over extended time periods are accounted for without the need for preprocessing.

The sum of all these advantages is segmentation under a wide range of magnifications, time resolutions, cell types and optical modalities that is both automated and robust. This results in the ability to track cells for hours or days and quantify a range morphological and phenotypic features without the need for user input, thus having broad applicability throughout live cell microscopy. The crux of the introduced self-supervised approach relies upon using the dynamic information embedded in each pixel – motion characterized via optical flow – as an elegant means to self-label cells versus background in time-lapse imagery. While cellular dynamics has long been appreciated as information rich with regards to understanding cell function, our approach demonstrates that it also provides the means for robust segmentation – a foundational step for achieving quantitative and objective live cell analysis.

Acknowledgements

The authors gratefully acknowledge the Devreotes laboratory of Johns Hopkins University for the *Dictyostelium discoideum* cell line. M.C.R. gratefully acknowledges support from the National Research Council Research Associateship Program and the Jerome and Isabella Karle Distinguished Scholar Fellowship Program. Funding for this project was provided by the Office of Naval Research through the Naval Research Laboratory's Basic Research Program and by the Biological Technology Office of the Defense Advanced Research Program Agency.

Author Contributions

Michael C. Robitaille: conceptualization, methodology, investigation, data curation, software, visualization, and writing. Jeff M. Byers: conceptualization, methodology, formal analysis, and software. Joseph A. Christodoulides: Resources, validation, and writing. Marc P. Raphael: conceptualization, funding acquisition, methodology, investigation, software, visualization, and writing.

Financial Conflicts of Interest

The authors do not have any conflict of interests with this work.

References

- 1 Caicedo, J. C., Singh, S. & Carpenter, A. E. Applications in image-based profiling of perturbations. *Current Opinion in Biotechnology* **39**, 134-142, doi:10.1016/j.copbio.2016.04.003 (2016).
- 2 Cadart, C., Zlotek-Zlotkiewicz, E., Le Berre, M., Piel, M. & Matthews, H. K. Exploring the Function of Cell Shape and Size during Mitosis. *Developmental Cell* **29**, 159-169, doi:10.1016/j.devcel.2014.04.009 (2014).
- 3 Zhou, X. B. & Wong, S. T. C. High content cellular imaging for drug development. *Ieee Signal Processing Magazine* **23**, 170-174, doi:10.1109/msp.2006.1598095 (2006).
- 4 Zhong, J. *et al.* Persistent hepatitis C virus infection in vitro: Coevolution of virus and host. *Journal of Virology* **80**, 11082-11093, doi:10.1128/jvi.01307-06 (2006).
- 5 Zhu, N. *et al.* Morphogenesis and cytopathic effect of SARS-CoV-2 infection in human airway epithelial cells. *Nature Communications* **11**, doi:10.1038/s41467-020-17796-z (2020).
- 6 Skylaki, S., Hilsenbeck, O. & Schroeder, T. Challenges in long-term imaging and quantification of single-cell dynamics. *Nature Biotechnology* **34**, 1137-1144, doi:10.1038/nbt.3713 (2016).
- 7 Caicedo, J. C. *et al.* Data-analysis strategies for image-based cell profiling. *Nature Methods* **14**, 849-863, doi:10.1038/nmeth.4397 (2017).
- 8 Deep learning gets scope time. *Nature Methods* **16**, 1195-1195, doi:10.1038/s41592-019-0670-x (2019).
- 9 Grys, B. T. *et al.* Machine learning and computer vision approaches for phenotypic profiling. *Journal of Cell Biology* **216**, 65-71, doi:10.1083/jcb.201610026 (2017).
- 10 Moen, E. *et al.* Deep learning for cellular image analysis. *Nature Methods* **16**, 1233-1246, doi:10.1038/s41592-019-0403-1 (2019).
- 11 Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology* **7**, doi:10.1186/gb-2006-7-10-r100 (2006).
- 12 Al-Kofahi, Y., Zaltsman, A., Graves, R., Marshall, W. & Rusu, M. A deep learning-based algorithm for 2-D cell segmentation in microscopy images. *Bmc Bioinformatics* **19**, doi:10.1186/s12859-018-2375-z (2018).
- 13 Falk, T. *et al.* U-Net: deep learning for cell counting, detection, and morphometry (vol 16, pg 67, 2019). *Nature Methods* **16**, 351-351, doi:10.1038/s41592-019-0356-4 (2019).
- 14 Sommer, C., Straehle, C., Kothe, U., Hamprecht, F. A. & Ieee. in *2011 8th Ieee International Symposium on Biomedical Imaging: From Nano to Macro IEEE International Symposium on Biomedical Imaging* 230-233 (2011).
- 15 Raphael, M. P., Sheehan, P. E. & Vora, G. J. A controlled trial for reproducibility. *Nature* **579**, 190-192, doi:10.1038/d41586-020-00672-7 (2020).
- 16 Beauchemin, S. S. & Barron, J. L. The computation of optical flow. *ACM Comput. Surv.* **27**, 433-467, doi:10.1145/212094.212141 (1995).
- 17 Farneback, G. in *Image Analysis, Proceedings Vol. 2749 Lecture Notes in Computer Science* (eds J. Bigun & T. Gustavsson) 363-370 (2003).
- 18 Robitaille, M. C., Byers, J. M., Christodoulides, J. A. & Raphael, M. P. Robust Optical Flow Algorithm for General, Label-free Cell Segmentation. *bioRxiv*, 2020.2010.2026.355958, doi:10.1101/2020.10.26.355958 (2020).

- 19 Schroeder, T. Long-term single-cell imaging of mammalian stem cells. *Nature Methods* **8**, S30-S35, doi:10.1038/nmeth.1577 (2011).
- 20 Jaccard, N. *et al.* Automated Method for the Rapid and Precise Estimation of Adherent Cell Culture Characteristics from Phase Contrast Microscopy Images. *Biotechnol. Bioeng.* **111**, 504-517, doi:10.1002/bit.25115 (2014).
- 21 Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F. & Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nature Methods* **15**, 917+, doi:10.1038/s41592-018-0111-2 (2018).
- 22 Vicar, T. *et al.* Cell segmentation methods for label-free contrast microscopy: review and comprehensive comparison. *Bmc Bioinformatics* **20**, 25, doi:10.1186/s12859-019-2880-8 (2019).
- 23 Wang, M. *et al.* Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy. *Bioinformatics* **24**, 94-101, doi:10.1093/bioinformatics/btm530 (2008).
- 24 Nath, S. K., Palaniappan, K. & Bunyak, F. in *Medical Image Computing and Computer-Assisted Intervention - Miccai 2006, Pt 1* Vol. 4190 *Lecture Notes in Computer Science* (eds R. Larsen, M. Nielsen, & J. Sporring) 101-108 (2006).
- 25 Buibas, M., Yu, D., Nizar, K. & Silva, G. A. Mapping the Spatiotemporal Dynamics of Calcium Signaling in Cellular Neural Networks Using Optical Flow. *Annals of Biomedical Engineering* **38**, 2520-2531, doi:10.1007/s10439-010-0005-7 (2010).
- 26 Delpiano, J. *et al.* Performance of optical flow techniques for motion analysis of fluorescent point signals in confocal microscopy. *Machine Vision and Applications* **23**, 675-689, doi:10.1007/s00138-011-0362-8 (2012).
- 27 Lee, R. M. *et al.* Quantifying topography-guided actin dynamics across scales using optical flow. *Mol. Biol. Cell* **31**, 1753-1764, doi:10.1091/mbc.E19-11-0614 (2020).
- 28 Meyers, J., Craig, J. & Odde, D. J. Potential for control of signaling pathways via cell size and shape. *Current Biology* **16**, 1685-1693, doi:10.1016/j.cub.2006.07.056 (2006).
- 29 Rangamani, P. *et al.* Decoding Information in Cell Shape. *Cell* **154**, 1356-1369, doi:10.1016/j.cell.2013.08.026 (2013).
- 30 Akanuma, T., Chen, C., Sato, T., Merks, R. M. H. & Sato, T. N. Memory of cell shape biases stochastic fate decision-making despite mitotic rounding. *Nature Communications* **7**, doi:10.1038/ncomms11963 (2016).
- 31 Robitaille, M. C. *et al.* Problem of Diminished cRGD Surface Activity and What Can Be Done about It. *Acs Applied Materials & Interfaces* **12**, 19337-19344, doi:10.1021/acsami.0c04340 (2020).