Check for updates

OPEN

Physics-based machine learning for subcellular segmentation in living cells

Arif Ahmed Sekh[1], Ida S. Opstad[1], Gustav Godtliebsen[2], Åsa Birna Birgisdottir[2,3], Balpreet Singh Ahluwalia[1], Krishna Agarwal[1] and Dilip K. Prasad[4 ✉]

Segmenting subcellular structures in living cells from fluorescence microscope images is a ground truth (GT)-deficient problem. The microscopes' three-dimensional blurring function, finite optical resolution due to light diffraction, finite pixel resolution and the complex morphological manifestations of the structures all contribute to GT-hardness. Unsupervised segmentation approaches are quite inaccurate. Therefore, manual segmentation relying on heuristics and experience remains the preferred approach. However, this process is tedious, given the countless structures present inside a single cell, and generating analytics across a large population of cells or performing advanced artificial intelligence tasks such as tracking are greatly limited. Here we bring modelling and deep learning to a nexus for solving this GT-hard problem, improving both the accuracy and speed of subcellular segmentation. We introduce a simulation-supervision approach empowered by physics-based GT, which presents two advantages. First, the physics-based GT resolves the GT-hardness. Second, computational modelling of all the relevant physical aspects assists the deep learning models in learning to compensate, to a great extent, for the limitations of physics and the instrument. We show extensive results on the segmentation of small vesicles and mitochondria in diverse and independent living- and fixed-cell datasets. We demonstrate the adaptability of the approach across diverse microscopes through transfer learning, and illustrate biologically relevant applications of automated analytics and motion analysis.

Machine learning is proving vital for cell-scale optical microscopy image and video analysis for studies in the life sciences[1–7], including major tasks such as cell segmentation, tracking, classification and population analytics. However, artificial intelligence and machine learning solutions are lacking for the analysis of small and dynamic subcellular structures such as mitochondria and vesicles. This affects how subcellular mechanisms are studied in life-science studies. First, only small-scale analyses can be conducted—via tedious manual annotations (Fig. 1a) and with limited conclusiveness owing to manual subjectivity (Fig. 1b) and the small statistical sample size. Second, because machine learning methods are rare in live-cell analysis because of the difficulties of annotation, other forms of quantitative statistical analyses such as fluorescence correlation spectroscopy are used instead. Third, electron microscopy image analysis is used for subcellular morphological investigations, but this does not provide the perspective of real-time unfolding of subcellular mechanisms. The information acquired using these techniques is valuable, but a computer-vision (CV) centric approach for observing live-cell subcellular processes holds an untapped potential for gaining unprecedented insights.

The segmentation of subcellular structures is a fundamental step towards realizing CV of subcellular mechanisms. Our interest lies in segmenting small and dynamic subcellular structures in cells from fluorescence microscope images. This is an immensely difficult task because of the small sizes of subcellular structures with respect to both the optical and digital resolutions of the microscopes. The structures often have dimensions on the order of 100–1,000 nm, while the pixel size in microscope images is generally 80–120 nm and the optical-resolution limit of advanced live-cell-compatible microscopes is typically 200–300 nm. This means that the details of the structures are often lost. Furthermore, the point spread function

(PSF) of optical microscopes induces a three-dimensional (3D) blur, as a result of which out-of-focus structures appear with different intensities and blurring profiles to those that are in focus. Accordingly, the segmentation of out-of-focus structures is often inaccurate. There are also other problems. Small structures have few binding locations for fluorophores, resulting in a low fluorescence intensity per structure. Small structures in living cells are highly dynamic and demand high-speed imaging (10–100 ms per frame), thus requiring short exposure times and limited fluorescence intensity. The signal-to-noise ratio is thus quite poor (2 to 4 in our experiments), which compounds the difficulty of segmentation. The great variability in the structures and the possibility of multiple overlapping structures creating high-intensity spots further complicate matters. These challenges are discussed in more detail in Supplementary Note 1 and Supplementary Fig. 1.

Image-processing techniques such as the popular Otsu approach[8] are consequently grossly inaccurate and contribute to large errors in conclusions about subcellular structures, as shown in the analytics results in Fig. 1a. Therefore, despite being fast (Fig. 1a), these techniques are not good candidates for performing subcellular analyses. Semi-supervised solutions also exist[9], but these are prone to subjectivity and tediousness, similar to manual segmentation. Supplementary Note 2 discusses the existing approaches in more detail, including manual annotation (Supplementary Fig. 2). Meanwhile, deep learning solutions hold promise for both live-cell analysis and large-scale systematic studies.

Interestingly, deep learning solutions optimized for cell-scale optical microscopy data[10] cannot simply be translated to subcellular scales because of the problems associated with digital and optical resolution and noise. In general, generating ground truth (GT) through manual segmentation over large datasets is considered the

[1]Department of Physics and Technology, UiT The Arctic University of Norway, Tromsø, Norway. [2]Department of Clinical Medicine, UiT The Arctic University of Norway, Tromsø, Norway. [3]University Hospital of North Norway, Tromsø, Norway. [4]Department of Computer Science, UiT The Arctic University of Norway, Tromsø, Norway. ✉e-mail: dilip.prasad@uit.no
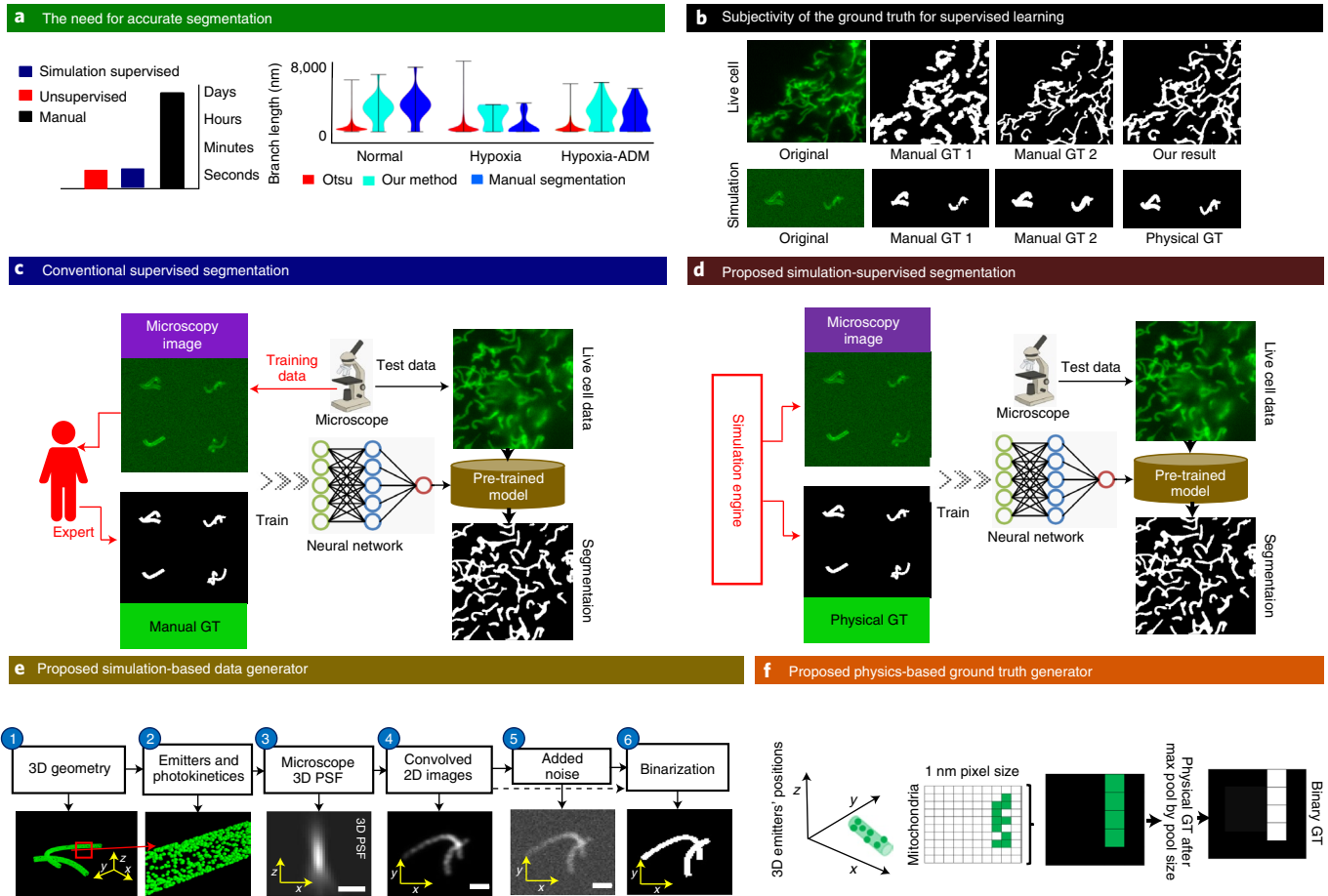
Fig. 1 | The motivation and approach for simulation-supervised deep learning. a, The time required for segmenting 30 frames from the LiveEpi1RatMitoRed dataset. Statistical analysis demands highly accurate segmentation (mitochondria branch length is shown as an example). Manual segmentation is the most reliable approach but is time consuming. Unsupervised methods such as Otsu generate inaccurate statistics due to inaccurate segmentation. Other methods suffer from different problems (Supplementary Note 2). ADM, adrenomedullin. **b**, Manual segmentation features variability related to the annotator's knowledge and expertise (manual GT 2 is an expert), so is unsuitable for supervised learning (Supplementary Note 1). **c**, The conventional approach employs manual segmentation. **d**, The proposed approach constructs a simulation-supervised training dataset with physics-based unambiguous GT. **e**, The proposed simulation-based data generation consists of six steps (illustrated for mitochondria, but the method is generalizable to any structure). (1) The 3D geometry of the structure is computed. (2) The locations and photokinetics of the fluorescent molecules are generated. (3) and (4) The microscope's 3D point spread function is applied to generate synthetic images. (5) Realistic noisy images are computed using a microscopy noise model. Scale bars, 1 μm. (6) The segmentation mask is computed from the noise-free image. The noisy image and the segmentation form one data pair in the simulation-supervised dataset. See Supplementary Notes 5–7 and Supplementary Fig. 7 for details of the simulation engine and the GT techniques. **f**, The proposed physics-based GT projects the locations of the fluorescent molecules on the lateral plane, then max-pools the projections by the microscope's pixel size, and finally performs binarization to form the GT. Neither the microscope nor the noise affect the GT.

only way to create training datasets. However, generating correct GT manually for fluorescent images of small subcellular structures is not possible, as the inaccuracy of every pixel contributes a non-negligible amount of error. Consequently, challenging subcellular structures such as mitochondria and vesicles have received less attention.

In this Article, we present a new physics-rooted deep learning approach for solving the GT deficiency in subcellular segmentation (Fig. 1). There are two key parts to our approach: (1) physics-based simulation-supervised learning, in which a supervised training dataset is created by simulating noisy microscope images, and (2) physics-based GT for generating the target segmentation in supervised learning. The simulation-supervised approach is a form of synthetic data-oriented approach. We present a short study about the ineffectiveness of synthetic data generation for our application in Supplementary Note 3 and Supplementary Fig. 3. Supplementary Note 4 presents a discussion on other known simulators for optical

microscopy with relevance to our problem and assesses the possibility of using super-resolution microscopy for generating synthetic microscopy datasets.

In our simulation-supervised approach, the training data are generated using a physics-based simulator that simulates everything from the binding location of an individual fluorescent molecule and its photokinetics to the 3D geometry of the subcellular structure on which fluorescent molecules are present, as well as the microscope instrument and noise characteristics. In addition, the physics-based simulation allows us to design a physics-based GT approach that is unbiased by the microscope instrument, free of manual subjectivity of segmentation, and assures that a particular geometry and fluorescent labelling result corresponds to a unique GT. This simulation engine and physics-based GT generates notably better and unbiased segmentation than expert-generated manual GT (Fig. 1b).

In this Article, we show that our simulation-supervision approach allows good-quality segmentation with a variety of deep
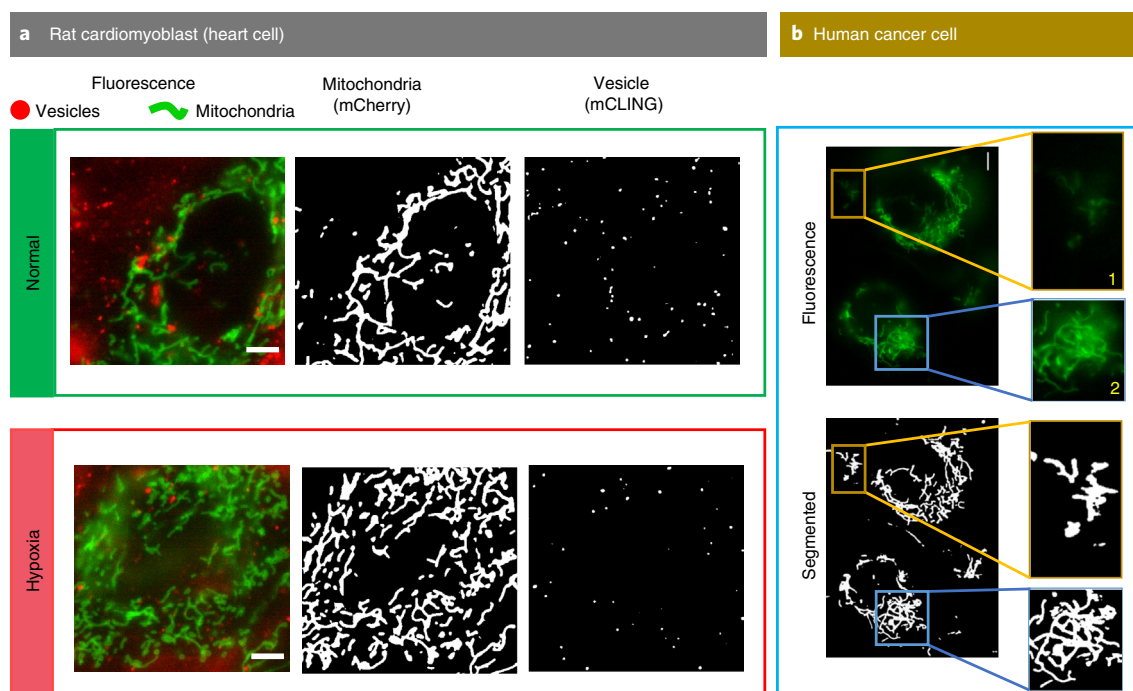
Fig. 2 | Examples of the segmentation of vesicles and mitochondria in living cells imaged using epifluorescence microscope Epi 1 and segmented using the proposed approach. a, Images of H9C2 cells (rat cardiomyoblasts) cultured under normal cell growth conditions as well as cells subjected to an hour of hypoxia up to several minutes before imaging. Changes in mitochondrial morphology are evident in the segmentation. Furthermore, although the membrane marker labels diverse membranous structures, the simulation-supervised deep learning model segments the vesicles and disregards the rest. **b**, The segmentation results for mitochondria in human cancer cells. The zoom-in of region 1 shows the ability of the proposed method to deal with low-intensity and low-contrast structures. The zoom-in of region 2 shows the ability to segment relatively dense mitochondrial structures while retaining details. Scale bars, 5 μm.

learning approaches, indicating its suitability for resolution-limited, noise-afflicted and GT-deficient subcellular microscopy data analysis. It can be applied across a variety of experimental conditions and cell types (Fig. 2) to segment organelles for which training was performed using simulation-supervised datasets (for example, mitochondria). We demonstrate the generalizability of our approach across microscopes through transfer learning (Fig. 3) and the possibility of performing multi-class classification, tracking and morphology-associated analytics at the scale of individual mitochondria (Figs. 4 and 5). We are also able to identify and analyse the interaction of mitochondria and vesicles inside living cells (Fig. 5).

Results on segmentation
We present the results of segmenting two types of subcellular entity—mitochondria and vesicles—in nine datasets of living and fixed cells (the datasets are described in Supplementary Fig. 4). Mitochondria and vesicles were chosen because they are interesting cases. Mitochondria are highly dynamic, tubular and shape-changing, with diameters close to the optical-resolution limit (200–300 nm) and lengths easily exceeding the depth of the focal field, rendering their segmentation challenging. Vesicles have simple geometries but vary significantly in size, with some being smaller than the optical-resolution limit and comparable to the digital resolution (pixel size of ~100 nm) and others being much larger. They thus present large variability in optical intensity and visibility. In fact, as our results indicate, vesicles are more challenging to segment than mitochondria, despite having a more simple geometry. Another point of interest is that subcellular mechanisms involving mitochondria and their interactions with vesicles such as endosomes and lysosomes are crucial for cell homeostasis and relevant for understanding disease development[11].

Physics-based simulation-supervised dataset and deep learning. Our simulation engine includes separate modules for the geometry of the subcellular organelles and labelling, the photokinetics of fluorescence, the microscope and image simulator, the noise simulator and the GT simulator. A detailed description of the six-step process (Fig. 1e) is presented in Supplementary Note 5 and Supplementary Fig. 5. This is extensible to further varieties of subcellular structures, microscopes and labelling protocols. At present it includes mitochondria and vesicle geometries (Supplementary Table 1), the ability to simulate epifluorescence and Airyscan microscopes (Supplementary Table 2), as well as surface labelling of vesicles and mitochondria. The simulation of photokinetics[12] is important for high-speed microscope videos (on the scale of milliseconds per frame) to model frame-to-frame variability. The modules are customizable to include other photokinetic and noise models. We created six simulation datasets, considering the two subcellular structures and three microscopes (Supplementary Table 3). Of these, datasets SimEpi1Mito and SimEpi1Vesi contain 7,000 images each. The other datasets contained 3,000 images each so as to explore the impact of the size of the simulation dataset on the accuracy of segmentation and the possibility of performing transfer learning across different microscopes.

We tested the efficacy of our simulation-supervised training as a suitable paradigm by using U-Net[13] (Supplementary Fig. 6) in conjunction with five state-of-the-art backbone networks. We found that Inception-V3 and EfficientNet-B3 generally performed best, although the other networks also performed robustly. For all further results we used the EfficientNet-B3 backbone. Details of the results are provided in Supplementary Note 6 and Supplementary Table 4. We also found that the performance is stable when the training dataset contains 3,000 or more simulation-supervised data samples (Supplementary Table 5).
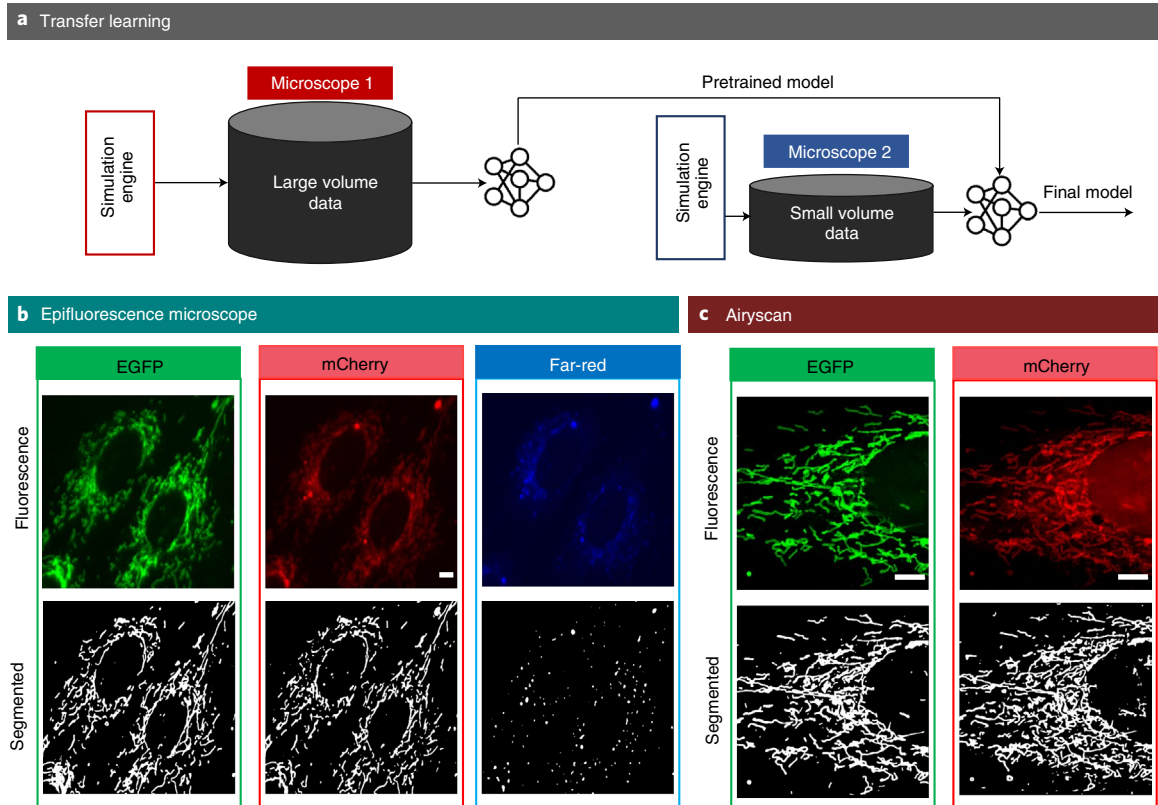
Fig. 3 | Transfer learning allows easy adaptation of our approach across different microscopes. a, A schematic depiction of transfer learning. Transfer learning involves using an existing deep learning model that has already been trained on a large simulation-supervised dataset from another optical fluorescence microscope and re-training this existing network using a small simulation-supervised dataset that has been custom created for the microscope of interest. **b**, Transfer learning can be done across two similar microscopes. A model pre-trained on simulation-supervised large data of microscope Epi 1 has been re-trained using small simulation-supervised data of microscope Epi 2 and applied on a living cell imaged using Epi 2 (datasets: LiveEpi2RatMitoRed, LiveEpi2RatMitoGreen and LiveEpi2RatVesiFarRed). Training across microscopes is sufficient; there is no need to train across dyes, as shown using tandem tagged mitochondria. **c**, Transfer learning is possible across different types of microscope as well. A model pre-trained on simulation-supervised large data of microscope Epi 1 has been re-trained using small simulation-supervised data of Airyscan microscope Airy 1 and applied on fixed cells imaged using Airy 1 (datasets: FixedAiry1RatMitoGreen and FixedAiry1RatMitoRed). Scale bars, 5 µm.

Physics-based GT. The physics-based simulator provides us with a unique opportunity to test different strategies for GT segmentation (Supplementary Note 7). Because the raw microscope images are generated before inclusion of the noise model, we explored the use of noise-free images with conventional morphological processing for generating the GT. We evaluated Otsu's thresholding[8] as well as Otsu's thresholding followed by morphological erosion by a kernel of the size of the microscope PSF to compensate for the blur of the microscope. A third technique was explored considering the role of noise in affecting how well an expert can segment images. For this, we thresholded using the noise level. Finally, we used the projection of the actual emitter distribution on the image plane directly as a mechanism of physics-based GT, as shown in Fig. 1f. This approach is not affected by the PSF or the noise level, and provides a unique unambiguous GT for a given sample. Supplementary Fig. 7 shows a comparison of these different methods for generating the GT.

We compared the four GT methods to identify which was the best strategy (Supplementary Note 7 and Supplementary Table 6). We found that the physics-based GT allows the deep learning models to perform better than the other GT mechanisms. Even a visual comparison of physics-based GT with a manual expert's segmentation (manual GT 2), as presented in Fig. 1b, presents a very good match. It also outperforms expert annotations in the challenging situations of out-of-focus structures and high noise levels (Supplementary Fig. 1c,d).

We also assessed the sensitivity of the performance of simulation-supervised training to important aspects of simulation (Supplementary Note 8). We note that our approach performs better if the simulation conditions closely match the experimental conditions (Supplementary Table 6 and Supplementary Fig. 8). The sensitivity of our approach contributes to good selectivity.

Comparison with contemporary techniques. We compared the performance of current methods used for segmenting subcellular structures from optical microscope images, namely (1) automated image-processing techniques such as Otsu-based thresholding[8], adaptive thresholding[14] and backpropagation[15], (2) semi-automatic segmentation techniques[9] and (3) the proposed simulation-supervised deep learning approach. The details of applying these methods are presented in the Methods. Table 1 presents the mean intersection over union (mIOU) values and the F1 scores. For the proposed method, we trained one deep learning model each for mitochondria and vesicles and used them with their corresponding test datasets. For the simulated test data of SimEpi1Mito and SimEpi1Vesi with physics-based GT, the proposed approach gives an advantage of ~10% for mitochondria and ~18% for vesicles.

We assessed whether the simulation-supervised training approach presents advantages over training a fresh and a pre-trained network with manually annotated GT (Supplementary Note 9 and Supplementary Table 7). Our results show that, even when manual
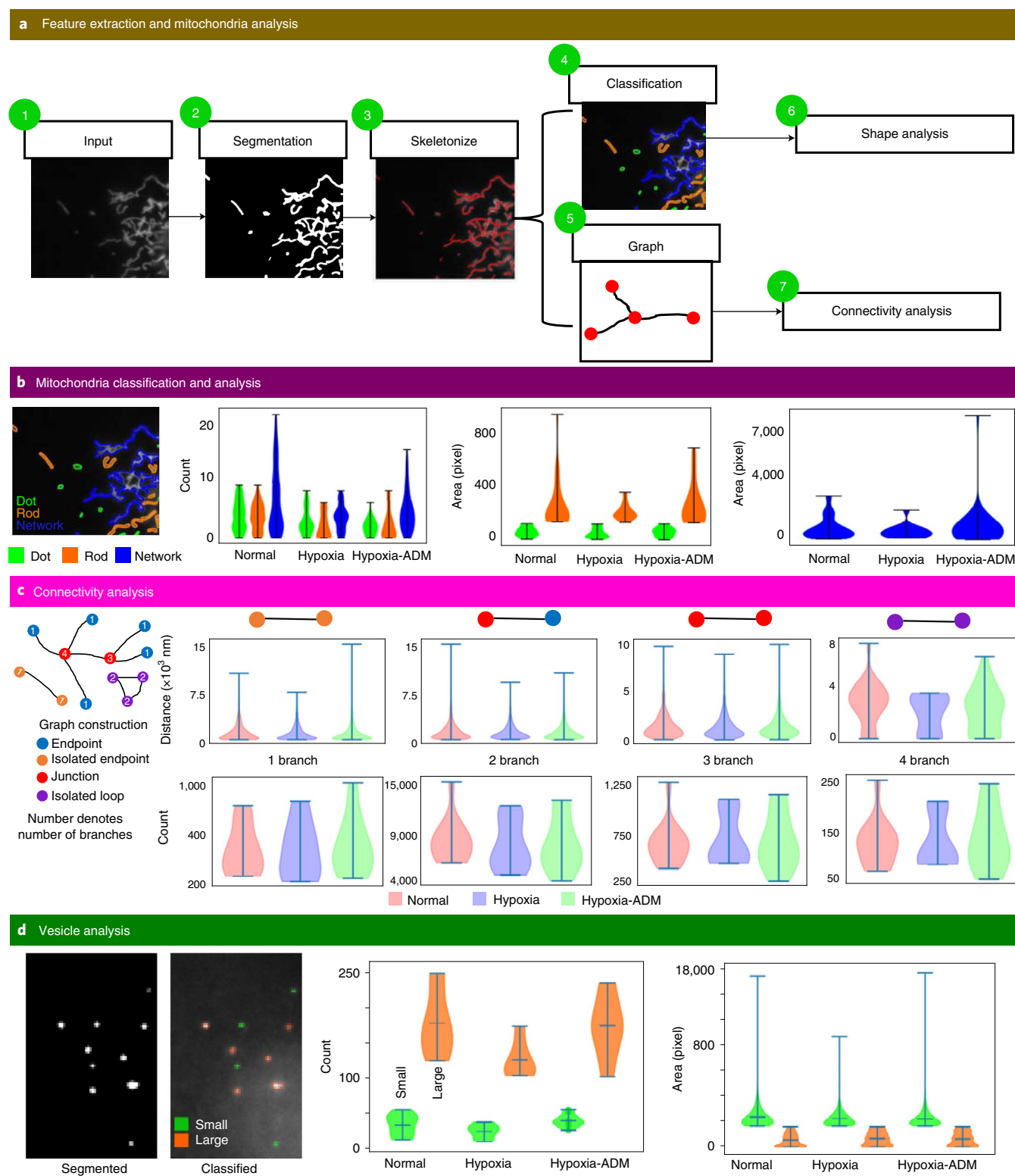
Fig. 4 | Automated segmentation applied to derive morphology-based analytics of mitochondria and vesicles. a, Steps for mitochondria analysis. The mitochondria are segmented and their skeletons are extracted. Based on the area and skeleton type, the mitochondria are classified in {dot, rod, network}. Each skeleton is converted to a graph and used for generating the analytics. **b**, Analytics derived from the morphologies for mitochondria ({dot, rod, network}), imaged under three different cell growth conditions. **c**, Graph-based connectivity analysis derived from the skeletons of mitochondria. **d**, Analytics derived from the morphologies of vesicles in the categories ({small, large}), imaged under three different cell growth conditions.

annotation is used as the GT, the proposed simulation-supervised training approach outperforms training using the manually annotated dataset. We also assessed whether it helps to use a larger

dataset with manual annotation generated by a sophisticated consensus-based annotation approach. We evaluated manual annotations by 12 scientists with the relevant background and obtained
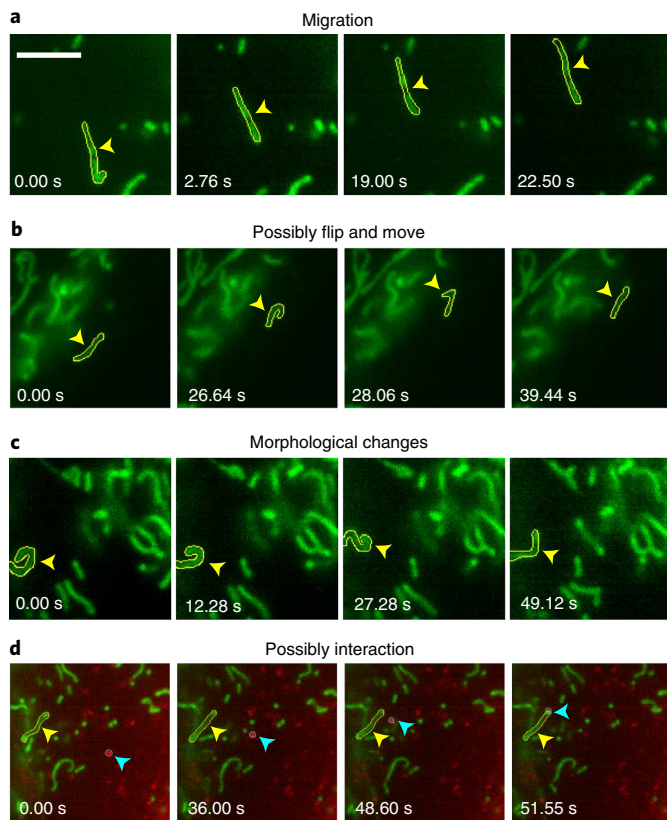
Fig. 5 | Temporal behaviour and events of mitochondria can be derived in an automated manner from segmentation. Four different behaviours manifested as morphological changes over time are shown. **a**, A mitochondrion moves from one place to another. **b**, A specific pattern of movement when a mitochondrion possibly flips and moves. **c**, Change in morphology over time. **d**, A possible interaction between a vesicle and a mitochondrion. Scale bar, 5 μm.

consensus on the GT. The consensus was quite unreliable, even with multiple annotators (Supplementary Fig. 2 and Supplementary Note 2). Also, the performance evaluation showed a large standard deviation as well as a large difference from the performance obtained with the physics-based GT (Supplementary Note 9 and Supplementary Table 8). Furthermore, the time taken by different annotators ranged from 3 to 10 min for annotating eight mitochondria in five small images. This task could thus prove considerably demanding in terms of resources if a large pool of annotators were employed for generating annotations for a sufficiently large dataset.

Live-cell segmentation results. We generated results on the live-cell datasets LiveEpi1RatMitoRed, LiveEpi1RatVesiFarRed and LiveEpi1HuMitoGreen. These datasets were acquired on epifluorescence microscope Epi 1 and manually annotated by an expert. Table 1 presents the mIOU values and F1 scores. The simulation-supervised deep learning method provides mIOU values of 74–76% and outperforms the closest method by 7–8% for mitochondria and ~13% for vesicles. A qualitative comparison of all the methods indicates that the proposed approach indeed provides the best results (Supplementary Fig. 9). Representative results for the proposed method (Fig. 2) illustrate that our approach can be applied to cells of different kinds and species (rat cardiomyoblasts are shown in Fig. 2a and human cancer cells in Fig. 2b). It can also be applied across cells subjected to different growth conditions, for example, under normal cell growth conditions (Fig. 2a, top row) and in cells subjected to hypoxia for 1 h before being imaged (Fig. 2a, bottom

row). Further illustrative examples indicate that there are situations where the expert annotation appears deficient in comparison to the CV-based segmentation result due to factors such as out-of-focus light and noise (Supplementary Figs. 10 and 11). Interestingly, in the LiveEpi1RatVesiFarRed dataset, a membrane marker is used that labels a huge variety of membrane structures including large structures and not just vesicles. Therefore, the raw microscope images of vesicles show a lot of other details beside vesicles. Nevertheless, the deep learning method shows clear proficiency at selecting vesicles, similar to trained experts. We also present interesting results in Fig. 2b that pertain to segmentation of mitochondria in human cancer cells from the LiveEpi1HuMitoGreen dataset. The zoomed view of region 1 in Fig. 2b shows the ability of the proposed method to tackle low-intensity regions. Conventional methods are unable to deal with such a situation where there are also other higher-intensity regions. In addition, dense mitochondrial regions and mitochondrial networks are a bigger challenge for expert annotation than the proposed segmentation approach, which yields good segmentation results that retain a lot of structural detail, for example, as shown in zoomed-in region 2 of Fig. 2b (also Supplementary Fig. 12). The quantitative performance of only 74−76% may thus be attributed to imperfect manual annotations as well.

Generalizability across microscopes and fluorophores. The generalizability across microscopes through computation-inexpensive re-training is crucial for quick adoption of this approach in various bioimaging laboratories across a variety of imaging set-ups. We thus assessed whether a simulation-supervised approach is amenable to transfer learning across microscopes (Fig. 3a). Details of the experiments and results are provided in Supplementary Note 10 and the results in Supplementary Table 9. We first considered two epifluorescence microscopes with different optical parameters (Supplementary Table 2). The results indicate a significant improvement in segmentation after transfer learning for challenging cases (Supplementary Fig. 13). We next considered transfer learning from an epifluorescence microscope to a different type of microscope, the Airyscan microscope. Here, as well, the results indicate a clear enhancement in the quality of segmentations and imply improved interpretability (Supplementary Fig. 14).

We also assessed the generalizability of our approach across fluorophores (Supplementary Note 8). We note that our approach performs robustly if the fluorophores used in experiments are different from the emission wavelength used for the simulation-supervised dataset. At the same time, we note that our approach presents high structural specificity using the challenging example presented in Supplementary Note 11 and Supplementary Fig. 15. Generalizability across fluorophores without re-training is a significant advantage because, when a deep learning model is trained for one subcellular structure and microscope, it can be used in a versatile manner for a wide range of biological experiments, irrespective of fluorophore, cell type and cellular conditions.

Application to morphological analysis
We demonstrate two applications: deriving morphological data analytics and event detection by tracking. This is made possible due to the better quality of segmentation over a large population of individual subcellular structures across optical microscope images of several cells under different conditions.

Morphology-based analytics. Morphological classification of mitochondria as dots, rods and network is highly informative[16]. The relevant statistics include the number and size of different mitochondrial phenotypes[17]. The primary challenge of such an analysis is accurate segmentation, which has been resolved. Figure 4a presents the steps of our analysis (also discussed in the Methods). Figure 4b presents the statistics of different morphologies and Fig. 4c depicts

Table 1 | Comparison of state-of-the-art subcellular segmentation methods with the proposed method

Type	Method	mIOU/F1 score				
		Simulation		Live cell		
		SimEpi1Mito	SimEpi1Vesi	(1)	(2)	(3)
Unsupervised	Otsu[8]	0.82/0.79	0.70/0.72	0.69/0.71	0.61/0.68	0.67/0.59
	Adaptive thresholding[14]	0.56/0.51	0.51/0.57	0.51/0.53	0.49/0.51	0.51/0.54
	Backpropagation[15]	0.68/0.67	0.55/0.58	0.58/0.61	0.51/0.54	0.54/0.57
Semi supervised	ImageJ plugin[9]	0.76/0.81	0.66/0.67	0.62/0.63	0.61/0.62	0.63/0.62
Simulation supervised	U-Net + EfficientNet-B3	0.92/0.94	0.89/0.91	0.76/0.84	0.74/0.79	0.75/0.82

We extensively analysed the segmentation performance of state-of-the-art unsupervised, semi-supervised and the proposed simulation-supervised methods. The experiment was conducted using the simulation-supervised datasets SimEpi1Mito and SimEpi1Vesi (as these were the base datasets and contained sufficient simulation data) and the experimental datasets (1) LiveEpi1RatMitoRed, (2) LiveEpi1RatVesiFarRed and (3) LiveEpi1HuMitoGreen. The mIOU values and F1 scores are presented side by side, separated by '/'.

graph-based connectivity analysis in cells subjected to three different cell growth conditions in the dataset LiveEpi1RatMitoRed. A similar analysis of small and large vesicular structures is presented in Fig. 4d. These results indicate the potential for (1) large-scale automated analysis under different cell growth conditions and (2) automated analysis of the evolution of cell health under dynamically changing cell-culture conditions. We present an additional analysis of the effect of carbonyl cyanide *m*-chlorophenyl hydrazone, a drug known to alter mitochondrial membrane potential, on mitochondrial morphologies observed in a living cell over a period of 60 min after administering the drug (Supplementary Note 12 and Supplementary Fig. 16).

Tracking and analysis of morphologically significant events. We assigned identities to each segmented mitochondrion and tracked them in high-speed microscopy videos of living cells. Because the segmentation both encodes morphologies and enables temporal tracking, the morphological changes over time can be monitored. We note different motion patterns that may be biologically relevant. Four examples are presented in Fig. 5 and one in Supplementary Fig. 17 (Supplementary Videos 1–5). Figure 5a shows a migrating mitochondrion. The mitochondrion in Fig. 5a was segmented by the expert as two mitochondria until the expert observed that the mitochondrion moves as a single entity. However, simulation-supervision segmented it as a single mitochondrion, compensating automatically for the out-of-focus region in the middle of it. In Fig. 5b, a mitochondrion performs a flip-and-move manoeuvre over ~40 s. Figure 5c shows a typical morphological change of a mitochondrion, from curled to elongated. Figure 5d shows an interesting situation where a vesicle migrates towards a mitochondrion in a seemingly targeted manner and then interacts with the mitochondrion. This analysis of such an event using CV demonstrates the utility of simulation-supervised deep learning-based segmentation for advanced analytics and analysis. The ability to perform automated detection of such events could lead to correlated behaviour analysis.

Discussion and conclusion

The proposed method brings physics and machine learning to a nexus where machine learning can create a significant impact with the help of physics-based modelling. Physics-based simulation-supervised training is thus proven to be the vital solution to the challenging GT-deficient problem of segmentation using deep learning for subcellular structures. The newly defined physics-based GT allows deep learning to tackle the optically hard problem of out-of-focus light and PSF-associated blurring. It also enables correct identification of the structures that the models have been trained to recognize, even in the challenging cases of fluorescence bleed-through (Supplementary Fig. 15). This approach is also generalizable across different types of cell and fluorophore. Transfer learning using smaller microscope-specific simulation-supervised datasets is a suitable mechanism for adopting the proposed paradigm across various fluorescence microscopy systems. Although the approach itself is generalizable, the models trained using this approach are sufficiently discriminative of the experimental conditions. Such discriminative ability is of significance in avoiding misleading inferences.

Valuable biological knowledge can be derived from automated, accurate segmentation of hundreds of subcellular organelles across one cell as well as in several cell images and long live-cell videos. Thus, several opportunities for performing advanced CV tasks are enabled by our proposed segmentation approach. Here, two proof-of-concept applications show the ability to perform morphology analysis and morphology-derived analytics. More application-specific morphological features of interest can be derived. Further morphological dynamics-based features may be analysed through tracking and following the morphological changes of segmented structures, for example, as shown in Fig. 5. The results strongly suggest that our approach is applicable to a wide range of different automated analysis pipelines. Accordingly, it may advance research in a variety of fields of biology and biomedicine in which the results and fundamental knowledge are often derived from bioimage analysis.

We highlight that the proposed method establishes the utility of a physics-based simulation-supervised training approach for deep learning applications in the microscopy data of living cells. This will open other research avenues in the future. More challenging and complicated structures of interest in the life sciences, such as the endoplasmic reticulum and Golgi bodies, can be simulated to extend the applicability of this approach in life-science studies. Furthermore, it will be interesting to explore whether 3D segmentation can be derived from raw microscopy image stacks with only a few *z*-planes, enabling long-term live-cell 3D analysis to be undertaken and more accurate observations to be derived (an example case is shown in Supplementary Fig. 17). Another important extension of this approach could be for label-free microscopy modalities such as bright-field microscopy. However, the realization of accurate physics-based simulation models for small structures will be a significant challenge[18], because the inherent optical contrast of the structures contributes to multiple scattering in the near field, which requires mathematically nonlinear physics solvers. Still, such models might be realized in the future and optimized for the large-scale creation of datasets of complex structures. We expect the nexus between machine learning and biology to only grow stronger, in the near future revolutionizing both our insights about biological systems and the opportunities available to researchers in the life sciences.

Methods

Physics-based simulation and GT mechanisms. The simulation flowchart is shown in Fig. 1e (further extended in Supplementary Fig. 5) and the simulation approach is presented in detail in Supplementary Note 5. The GT mechanisms are presented in Fig. 1f and detailed in Supplementary Note 7. The simulation was implemented on a Windows-based computer combined with Python 3.6. The simulator is shared for public use (see 'Code availability' section).

Preparation of the simulation data for training and testing. We used three different microscopes for imaging and therefore use similar configurations for the simulations to individually create the training datasets for each microscope (Supplementary Table 3). Our simulation framework can generate 128 × 128 image pairs (image and segmentation GT). We combine four such independent images to create a 2 × 2 tile with dimensions of 256 × 256. We use two types of training batch. The first batch is a large volume of data (7,000 tiles) of a specific microscope used for baseline experiments and the deep model is trained from scratch. The second batch is generated for two different microscopy settings and a comparatively small amount of data (3,000 slides). This batch is used to find the effect of transfer learning. We use standard data augmentation such as flip, rotation and so on during training. We consider 60% training, 20% validation and 20% for testing for each simulation dataset. The method is repeated for two different subcellular structures: mitochondria and vesicular structures.

U-Net backbones and training details. The backbones used in the U-Net encoder for Supplementary Table 4 are ResNet50, ResNet 100[19], VGG16[20], Inception[21] and EfficientNet-B3[22]. In our research, we found EfficientNet-B3 as the best-performing encoder in terms of validation accuracy on the simulation-supervised mitochondria dataset. This consists of three mobile inverted bottleneck convolution layers integrated in the encoder. For each network, the input and output are 256 × 256. We use standard data augmentation such as flip, rotation and so on where applicable. Early stopping and learning rate reduction are also used based on the mIOU. All experiments are carried out using an Intel(R) Xeon(R) Gold 6154 CPU with 128 GB of RAM and an NVIDIA Quadro RTX 6000 GPU with capacity of 24 GB.

Contemporary methods for subcellular segmentation. Here, we present details of the implementation of the contemporary methods (the results are presented in Table 1 and Supplementary Fig. 9). Otsu-based thresholding[8] and adaptive thresholding[14] use a histogram of the intensities for segmentation. These are non-parametric methods and therefore do not require any user input. We use the OpenCV 3.4 library combined with Python 3.7 for the implementation. The ImageJ-based morphological plugin[9] (MorphoLibJ v1.4.1) is used on the Windows platform with default parameter settings. Manual thresholding is implemented using the OpenCV 3.4 library combined with Python 3.7, and a suitable threshold for best-performing segmentation is extracted by varying the global threshold over the intensity histogram of the images. The iterative backpropagation-based segmentation[15] is implemented in the OpenCV 3.4 library combined with Python 3.7, and we use 1,000 iterations for the segmentation benchmarking.

Evaluation metric. The metric used for quantification of the performance of segmentation is mIOU. This is a state-of-the-art metric used in segmentation problems[10]. mIOU values are calculated by taking the ratio of the overlapped segmented area and the union of the segmented area between the GT and segmented image, that is $TP/(TP + FP + FN)$, where true-positive (TP), false-positive (FP) and false-negative (FN) regions are used.

Processing related to the morphological analysis presented in the main text. First, we apply our simulation-supervised deep learning model for segmentation. Next, a rule-based classification on the segmented area of individual mitochondria is used to classify the rod, dot and network morphologies. The mitochondria are classified into three categories: dot, rod and network. First, the mitochondria are segmented using the proposed method. Next, the binary images are converted into a skeleton and a graph is constructed according to ref. [23], where nodes are endpoints or mitochondria junctions. The degree of a node (d) is the number of branches connected to the node. Finally, each graph is classified as dot, rod and network using

$$\text{Class}_{\text{mitochondria}} = \begin{cases} \text{Dot,} & \text{if area} \leq 120 \text{ pixel and } \max(d) = 1 \\ \text{Rod,} & \text{if area} > 120 \text{ pixel and } \max(d) = 1 \\ \text{Network,} & \text{otherwise} \end{cases}$$

The statistics of the frequency of occurrences of different types of mitochondria (dot, rod and network) and area are presented in Fig. 4b as violin plots. There are 30 cells in our live-cell dataset and the mean and standard deviation are calculated for each cell. In a similar manner, the vesicles are classified into two categories: large and small. The segmented vesicles are fitted inside circles. The vesicles are classified using a heuristic threshold of the radius (r) as

$$\text{Class}_{\text{vesicle}} = \begin{cases} \text{Small,} & \text{if } r \leq 150 \text{ nm} \\ \text{Large,} & \text{otherwise} \end{cases}$$

A complex graph-based connectivity analysis is also explored by converting the segmented images into skeletons and graphs. After obtaining the graphs, the nodes are classified as shown in Fig. 4c. If a graph contains junction nodes, it is a network. The analytics show more nuanced information about networks through the endpoint–junction lengths and junction–junction lengths.

Indeed, the classification can be performed using simple rules such as used here, or fuzzy rules, more elaborate rules and even deep learning approaches may be employed for morphological classification depending on the need of the applications.

Tracking of mitochondria and vesicles. First, the proposed U-Net-based segmentation is used to segment the subcellular structures. Then, the Kalman filter and Hungarian algorithm[24] are employed to track individual structures over time[25].

Microscopes and imaging parameters. Three different microscopes were used in this work (Supplementary Table 2). The first, microscope Epi 1, is a GE DeltaVision Elite microscope and was used for datasets LiveEpi1RatMitoRed and LiveEpi1RatVesiFarRed. The exposure time for imaging the vesicles and mitochondria was 10 ms. The acquisition rate was 50 frames per second. The acquisition was performed in sequential mode. The LiveEpi1HuMitoGreen dataset was also recorded with this microscope. The second, microscope Epi 2, is a Zeiss CellDiscoverer 7 with a Plan-Apochromat ×50 water objective and an NA of 1.2. The LiveEpi2RatMitoGreen, LiveEpi2RatMitoRed and LiveEpi2RatVesiFarRed datasets were recorded with this microscope. The third, microscope Airy 1, is a Zeiss LSM 880 ELYRA with a C Plan-Apochromat ×63 oil objective with an NA of 1.4. The FixedAiry1MitoGreen, FixedAiry1MitoRed and FixedAiry1RatVesiBlue datasets were recorded using this microscope.

Cell culture and imaging conditions for the live-cell datasets LiveEpi1RatMitoRed and LiveEpi1RatVesiFarRed. The rat cardiomyoblast cell-line H9c2 (cells derived from embryonic heart tissue; Sigma-Aldrich) were cultured in high-glucose (4.5 g l⁻¹) Dulbecco's modified Eagle medium (DMEM) with 10% fetal bovine serum (FBS). The cells were transiently transfected using TransIT-LT1 (Mirus) to express the mitochondrial fluorescence marker mCherry-OMP25-TM (emission maximum at 610 nm). After 24 h of transfection, the cells were incubated in serum-free DMEM medium for 4 h and then the medium was changed back to DMEM with 2% serum just before treatment for 1 h (see below). After treatment, the medium was changed back to DMEM 10% FBS. The cells were divided into three pools: normal, hypoxia and hypoxia-ADM. For the normal conditions (control) pool, the cells were kept under normal cell growth conditions at 37 °C with about 21% O_2 and 5% CO_2. For the hypoxia pool, the cells were subjected to hypoxia (deficiency of oxygen; 0.3% O_2 level) by incubation in a hypoxic cell incubator for 60 min. For the hypoxia and ADM pool, the cells were subjected to hypoxia as for the cells above, but were simultaneously treated with the peptide hormone adrenomedullin (ADM) at a concentration of 10^{-6} M. This hormone has been found to exhibit protective functions under various pathological conditions, such as ischaemia in heart cells during myocardial infarction. The cells were labelled using the live-cell-friendly fluorescent marker mCLING-ATTO647N immediately before imaging using a concentration of 1:2,000 with a 12-min incubation time. After incubation, the medium was replaced with cell-culture medium (DMEM 10% FBS) for time-lapse microscopy at 37 °C, atmospheric oxygen (that is, the cells in hypoxia and hypoxia-ADM pools were no longer in an oxygen-deficient condition) and 5% CO_2. The membrane marker was quickly internalized by the cells and labelled small membrane-bound vesicles in the cells. This membrane marker exhibits a fluorescence emission maximum at a wavelength of 662 nm. The mitochondrial marker mCherry-OMP25-TN and membrane marker mCLING-ATTO647N were imaged using epifluorescence microscope Epi 1 sequentially in separate colour channels.

Cell culture and imaging conditions for the live-cell dataset LiveEpi1HuMitoGreen. MCC13 cells were maintained in an incubator at 37 °C with 20% O_2 and 5% CO_2, with a growth medium consisting of RPMI 1640 (Sigma-Aldrich) supplemented with 10% FBS (Sigma-Aldrich) and 1% penicillin/streptomycin (Sigma-Aldrich). The cultures used for experiments were thawed from stocks stored in liquid nitrogen a minimum of one week before labelling and imaging.

Labelling with CellLight Mitochondria-RFP BacMam 2.0 (Thermo Fisher Scientific) was carried out according to the manufacturer's protocol with 15 to 45 particles per cell (PPC) ~20 h before imaging. Transduced cells were grown under the same cell growth conditions as described above but in antibiotic-free medium.

Immediately before imaging, the cells were incubated with MitoTracker Deep Red (Thermo Fisher Scientific) for 30 min, then washed in phosphate-buffered saline or live-cell imaging medium.

These cells were imaged without the use of a microscope incubation system at room temperature (~25 °C), but in pre-heated (37 °C) live-cell imaging solution (Thermo Fisher Scientific).

Cell culture and imaging conditions for the live-cell datasets LiveEpi2RatMitoGreen, LiveEpi2RatMitoRed and LiveEpi2RatVesiFarRed. The rat cardiomyoblast cell-line H9c2 (described above), genetically modified using retrovirus to have stable expression of tandem tagged (mCherry-EGFP) mitochondrial outer membrane protein 25 (OMP25)-transmembrane domain (TM), was utilized. Equal expression of fluorescence intensity in cells was achieved through flow cytometry sorting. The cells were cultured in high-glucose DMEM with 10% FBS or in medium for glucose deprivation and galactose adaption. The glucose deprivation medium consisted of DMEM without glucose (11966-025, Gibco) supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 10 mM galactose, 10% FBS, 1% streptomycin/penicillin and 1 µg ml^{-1} of puromycin (InvivoGen, ant-pr-1). The cells were adapted to galactose for a minimum of seven days before experiments. The cells were seeded on MatTek dishes (P35G-1.5-14-C, MatTek Corporation) and imaged when they reached ~80% confluency. Labelling of lysosomes (acidic endosomal system) was carried out by treating cells for 30 min with 50 nM Lysotracker Deep Red (cat. no. L12492, Thermo Fisher) according to the manufacturer's recommendation. After labelling, the medium was replaced with fresh medium (described above) for live-cell microscopy. The cells were imaged at 37 °C with atmospheric oxygen and 5% CO_2. Imaging was performed using the Epi 2 microscope. For the live-cell imaging, selected positions were imaged for a duration of 10 min each, with one frame being taken every 5 s, giving 120 frames. Each frame consisted of a seven-slice z-stack with a 0.31-µm interval between slices.

Cell culture and imaging conditions for the fixed-cell datasets FixedAiry1MitoGreen, FixedAiry1MitoRed and FixedAiry1RatVesiBlue. mCherry-EGFP-OMP-25TM H9c2 cells were seeded on glass coverslips (1.5). The cell growth conditions and labelling were the same as used for the live-cell datasets in the previous paragraph (datasets 4, 5 and 6 in Supplementary Fig. 4. After labelling, the coverslips were washed once in phosphate-buffered saline, then fixed using 4% paraformaldehyde and 0.2% glutaraldehyde for 20 min at 37 °C. The coverslips were then washed in phosphate-buffered saline and mounted using Prolong glass antifade mountant (cat. no. P36980, Thermo Fisher). Imaging was performed using the Airy 1 microscope. Airyscan images were taken of regions of interest. All Airyscan images were processed using the LSM 880 Zen software packages 'Airyscan Processing' method.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The data used for the main analyses presented in the paper are available at https://doi.org/10.18710/IDCUCI[26].

Code availability
The code related to this work is available at https://doi.org/10.5281/zenodo.5017066[27].

References
1. Chang, S. W. & Liao, S. W. KUnet: microscopy image segmentation with deep Unet based convolutional networks. In *Proc. IEEE International Conference on Systems, Man and Cybernetics* 3561–3566 (IEEE, 2019).
2. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. & Liang, J. UNet++: a nested U-Net architecture for medical image segmentation. *Feep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support* **11045**, 3–11 (2018).
3. Van Valen, D. A. et al. Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLoS Comput. Biol.* **12**, e1005177 (2016).
4. Sadanandan, S. K., Ranefall, P., Le Guyader, S. & Wählby, C. Automated training of deep convolutional neural networks for cell segmentation. *Sci. Rep.* **7**, 7860 (2017).
5. Aydin, A. S., Dubey, A., Dovrat, D., Aharoni, A. & Shilkrot, R. CNN based yeast cell segmentation in multi-modal fluorescent microscopy data. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops* 753–759 (IEEE, 2017).
6. Majurski, M. et al. Cell image segmentation using generative adversarial networks, transfer learning and augmentations. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops* 1114–1122 (IEEE, 2019).
7. Sekh, A. A. et al. Learning nanoscale motion patterns of vesicles in living cells. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 14014–14023 (IEEE, 2020).
8. Xu, X., Xu, S., Jin, L. & Song, E. Characteristic analysis of Otsu threshold and its applications. *Pattern Recogn. Lett.* **32**, 956–961 (2011).
9. Arganda-Carreras, I. & Legland, D. Morphological segmentation (IJPB-plugins). *ImageJ* https://imagej.net/Morphological_Segmentation (2019).
10. Shibuya, E. & Hotta, K. Feedback U-Net for cell image segmentation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops* 974–975 (IEEE, 2020).
11. Chen, Q. et al. Quantitative analysis of interactive behavior of mitochondria and lysosomes using structured illumination microscopy. *Biomaterials* **250**, 120059 (2020).
12. Agarwal, K. & Macháň, R. Multiple signal classification algorithm for super-resolution fluorescence microscopy. *Nat. Commun.* **7**, 13752 (2016).
13. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention* (eds Navab N. et al.) 234–241 (Springer, 2015).
14. Shen, S. P. et al. Automatic cell segmentation by adaptive thresholding (ACSAT) for large-scale calcium imaging datasets. *eNeuro* **5**, 56 (2018).
15. Kanezaki, A. Unsupervised image segmentation by backpropagation. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing* 1543–1547 (IEEE, 2018).
16. Li, Y. et al. Imaging of macrophage mitochondria dynamics in vivo reveals cellular activation phenotype for diagnosis. *Theranostics* **10**, 2897–2917 (2020).
17. Leonard, A. P. et al. Quantitative analysis of mitochondrial morphology and membrane potential in living cells using high-content imaging, machine learning and morphological binning. *Biochim. Biophys. Acta* **1853**, 348–360 (2015).
18. Svoboda, D. & Ulman, V. Generation of synthetic image datasets for time-lapse fluorescence microscopy. In *Proc. International Conference Image Analysis and Recognition* (eds Campilho A. & Kamel M.) 473–482 (Springer, 2012).
19. Wu, Z., Shen, C. & Van Den Hengel, A. Wider or deeper: revisiting the ResNet model for visual recognition. *Pattern Recognit.* **90**, 119–133 (2019).
20. Qassim, H., Verma, A. & Feinzimer, D. Compressed residual-VGG16 CNN model for big data places image recognition. In *Proc. IEEE Annual Computing and Communication Workshop and Conference (CCWC)* 169–175 (IEEE, 2018).
21. Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. Inception-v4, inception-ResNet and the impact of residual connections on learning. In *Proc. 31st AAAI Conference on Artificial Intelligence* 4278–4284 (ACM, 2017).
22. Tan, M., Pang, R. & Le, Q. V. EfficientDet: scalable and efficient object detection. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 10781–10790 (IEEE, 2020).
23. Nunez-Iglesias, J., Blanch, A. J., Looker, O., Dixon, M. W. & Tilley, L. A new Python library to analyse skeleton images confirms malaria parasite remodelling of the red blood cell membrane skeleton. *PeerJ* **6**, e4312 (2018).
24. Weng, S.-K., Kuo, C.-M. & Tu, S.-K. Video object tracking using adaptive Kalman filter. *J. Vis. Commun. Image Represent.* **17**, 1190–1208 (2006).
25. Wojke, N., Bewley, A. & Paulus, D. Simple online and realtime tracking with a deep association metric. In *Proc. IEEE International Conference on Image Processing (ICIP)* 3645–3649 (IEEE, 2017).
26. Opstad, I. S., Godtliebsen, G. & Sekh, A. A. Replication data for: physics based machine learning for sub-cellular segmentation in living cells. *DataverseNO* https://doi.org/10.18710/IDCUCI (2021).
27. Sekh, A. A. skarifahmed/physeg: physics based machine learning for sub-cellular segmentation in living cells. *Zenodo* https://doi.org/10.5281/zenodo.5017066 (2021).

Author contributions
A.A.S., K.A. and D.K.P. conceptualized the project. A.A.S. and D.K.P. were responsible for the methodology and formal analysis. A.A.S. developed the software. K.A., D.K.P.,

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42256-021-00420-0.

Correspondence and requests for materials should be addressed to Dilip K. Prasad.

Peer review information *Nature Machine Intelligence* thanks Mark A. Anastasio and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Corresponding author(s): Dilip K. Prasad

Last updated by author(s): Apr 8, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	In the conventional sense, no software was used for data collection. However, simulated datasets were created using custom codes on Python platform. The details will be made available with the codes shared with the editor and reviewers according to the software policy.
Data analysis	In the conventional sense, data analysis for generating box plots have been done using custom codes in Python. Deep learning codes have been generated in Python on Keras. Codes of all these will be made available to the editor and reviewers according to the software policy.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Accession codes will be made available before publication. All the raw microscopy data used in this paper will be made open access. There are no restrictions on data availability in our current knowledge.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	In our case, the conventional sample size pertaining life sciences does not apply. Our sample size is typically determined by the number of images in the training or test datasets. These have been explicitly included in the paper.
Data exclusions	No data was excluded from the analysis
Replication	In the sense of conventional life sciences jargon, replication does not apply. Replication in the sense of deep learning applies in the conventional sense and has been verified through cross-validation.
Randomization	In the sense of conventional life sciences jargon, randomization does not apply. Randomization as relevant to machine learning has been incorporated using standard mechanisms.
Blinding	Our data does not require blinding of the investigators.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	H9c2 (2-1) Rat DB1X heart myoblast; Catalogue No: 88092904 ; Lot No. 08G008 P+8 ; Source: ECACC; European Collection of Authenticated Cell Cultures; The MCC13 cell line was a gift from Professor Ugo Moens, Molecular Inflammation Research Group, UiT, and originated from Professor Baki Akguel, University of Cologne, Institute of Virology, Germany.
Authentication	H9c2 (2-1) cell line was authenticated by DNA fingerprinting and quality controlled by source (ECACC) as stated in the Certificate of analysis for lot no 08G008; The MCC13 cells line was not authenticated.
Mycoplasma contamination	H9c2 (2-1) cells tested negative for mycoplasma; the MCC13 cell line was not tested for mycoplasma.
Commonly misidentified lines (See ICLAC register)	Not relevant for any cell line used.