

Fast imaging of 15 intracellular compartments and interactions by deep learning segmentation of super-resolution data

Karl Zhanghao

Southern University of Science and Technology <https://orcid.org/0000-0002-8331-6240>

Meiqi Li

Peking University <https://orcid.org/0000-0003-3586-176X>

Xingye Chen

Tsinghua University

Wenhui Liu

Tsinghua University

Tianling Li

China Agricultural University

Yiming Wang

Peking University

Fei Su

Southern University of Science and Technology

Zihan Wu

Southern University of Science and Technology

Chunyan Shan

College of Life Sciences, Peking University <https://orcid.org/0000-0002-5532-2811>

Jiamin Wu

Tsinghua University <https://orcid.org/0000-0003-3479-1026>

Yan Zhang

Peking University

Jingyan Fu

China Agricultural University <https://orcid.org/0000-0002-5896-8132>

Peng Xi

Department of Biomedical Engineering, College of Engineering, Peking University, Beijing 100871, China.

<https://orcid.org/0000-0001-6626-4840>

Dayong Jin (✉ dayong.jin@uts.edu.au)

University of Technology Sydney <https://orcid.org/0000-0003-1046-2666>

Keywords: deep learning, super-resolution, live-cell imaging

Posted Date: May 31st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1686675/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The number of colors in fluorescence microscopy is far less than the types of intracellular compartments, bringing gaps in studying live-cell anatomy and multiple organelles' interactions. Here, we report that super-resolution imaging in association with deep convolutional neuronal networks can predict 15 subcellular structures at > 91.7% pixel accuracy using one laser excitation and two detection channels. It not only bypasses the limitations of multi-color imaging with single dye labeling but also accelerates the imaging speed by more than one order of magnitude. We find that the super-resolution ratiometric images well reflect the heterogeneity of organelles as the intrinsic "optical fingerprint" and the neuronal networks can be generalized with transfer learning to predict both 3D and 2D datasets from different microscopes, different cell types, and even complexed system of living tissues. It enables us to resolve the 3D anatomic structure of live cells at different mitotic phases and to track down the fast dynamic interactions among 9 intracellular compartments.

Introduction

Until now, up to six fluorescent colors with multiple excitations and detections can be reported to simultaneously observe the 6 types of organelles' interactions and coordinations^{1,2}. Further increase in color will be intrinsically limited by the dyes' crosstalk in the spectrum domain. The duration, as the result of the multiple excitations and image acquisition steps, will proportionally increase in multi-color fluorescence imaging, and consequently, reduce the imaging speed. The phototoxicity, brought by the repetitive excitations, further becomes a critical issue in live-cell imaging.

The strategy of multicolor labeling of the sophisticated intracellular components can be tedious, non-specific, and inefficient. For organic dyes, many organelles are surrounded by lipid membranes that share similar surface and chemical properties, therefore, it is difficult to specifically label each organelle type. Though fluorescent proteins can achieve better specificity to label some types of organelles, the plasmids suffer from no expression or overexpression (Fig. S1a-f). The joint labeling efficiency will drop exponentially with the increase of the number of multiple probes. For example, the efficiency in co-expressing five fluorescence proteins will drop to only ~ 35%, assuming the expression efficiency of individual fluorophores should be at least 80% (Fig. S1g). Either the diffusion of organic dyes or the expression of fluorescent protein is much more difficult in living tissues, which causes many troubles in fluorescent labeling of living tissues (Fig. S1h, i).

With the emergence of artificial intelligence, deep learning has been recently introduced to the in-silico prediction of multi-color images from transmitted bright field images³⁻⁶, in which deep convolutional neuronal networks (DCNN) are trained to learn the relation between image pairs of bright field images and fluorescent images, and a set of fluorescent images can be predicted from the label-free images. Nevertheless, as the training images suffer from low resolution and low contrast, it is yet to be satisfactory in the prediction of subcellular structures. For organelles, including Golgi apparatus, with smaller size or low contrast, the Pearson Correlation Coefficient (PCC) of predicted images is only less

than 0.2, compared with the ground truth³. Although the predicted images of mitochondria show a high PCC, compared with the ground truth images, the valuable features of ultrafine structures cannot be resolved due to the diffraction limit. Moreover, as the 3D stacks of the bright field images are required to feed the training network, the long acquisition time makes it hard to capture the fast dynamics of the intracellular organelles. A recent work⁷ segments whole-cell organelles on electron microscopy data, which was also performed on fixed cells.

Super-resolution microscopy, by taking advantage of their high resolution and contrast, can image the subcellular structures and dynamics^{8–11}. While long-term imaging in live cells has been demonstrated^{10–12}, only two or three colors can be achieved. Besides the problem of the spectrum crosstalk, this is limited by the issues of photobleaching and phototoxicity, as to achieve both high spatial resolution and multicolor imaging intrinsically requires a large photon budget. Although deep learning has been utilized to enhance the signal-to-noise ratio, to reconstruct images, and to increase the spatial resolution, in-silico multicolor imaging in super-resolution microscopes remains unexplored.

Here, our strategy is to stain multiple intracellular compartments with one lipid dye that can stain the membranous structures within the live cell with nearly 100% efficiency. We further use a commercial super-resolution spinning disk microscope¹³ to obtain the super-resolution images (Fig. S2) and explore the organelle segmentation of the two-color super-resolution images with DCNN. Moreover, as the emission spectrum of the dye responds to the lipid polarity of membranes¹⁴, we further apply the dual-color detection to discriminate vesicle organelles with similar shapes and sizes. We demonstrate such a simple cell staining followed by the rapid acquisition of spatial (3D-SR) and spectral (2-channel splitting) imaging data can empower DCNN networks to predict 15 intracellular structures with high accuracy and throughput, opening a new paradigm for multiplexing imaging inside the living cell and multi-cellular system (Fig. 1).

Results

Data acquisition and preparation of the super-resolution ratiometric images

The live cells are stained by Nile Red and imaged by a commercial super-resolution microscope with doubled spatial resolution of the diffraction limit (Fig. S2). One laser (488 nm) excites the dye, and two emission channels (yellow: 580–653 nm, red: 665–705 nm) detect the fluorescent intensity and spectral ratio (microscope M1). The intensity images, as the result of averaging the yellow and red channels of Nile Red fluorescence images, and the ratio-metric images, by dividing the red-channel image with the yellow-channel one, are used as the input of the DCNN networks. Once the networks are trained, different types of organelles can be segmented directly from the super-resolution ratiometric images (Fig. 1a, Fig. S3). As the spectral ratio of Nile Red shows a heterogeneous distribution across different organelles, once the concentration of Nile Red set unchanged, the spectral ratio becomes rather consistent under different

excitation powers or in different cell types (Fig. S4). More encouragingly, we further find that the spectral ratio readouts of Nile Red on different organelles remain little changed even the images of cells are collected on a different microscope M2 whose excitation laser has been changed to 470 nm with the same detection bands (yellow: 580–653 nm; red: 665–705 nm). These results conclude that the spectral ratio serves as an intrinsic “optical fingerprint” for reliable and reproducible segmentation of organelle types.

To obtain the ground truth masks of different organelles, we acquire colocalization images using GFP plasmids that each time specifically label a type of organelle. To acquire the training dataset, the cells labeled by both the organelle-specific GFP and Nile Red are imaged with 488-nm excitation and three detection channels including the green emission channel (500–550 nm) for GFP. Fig. S5 confirms that the spectral ratio of Nile Red is not affected by the slight emission spectrum overlaps of GFP and Nile Red.

To obtain the binary masks, the foreground area needs to be segmented from the background (out-of-focus or dark) area in colocalization images. With conventional thresholding-based segmentation, different thresholding methods are required for different intracellular structures, with subject adjustment of parameters. When the cells in the same image have different expression levels of fluorescent proteins, their fluorescent intensities will vary so that local thresholding methods are required, which introduces more adjustable parameters. To solve this problem, an FG-net (Fig. S6b) was trained in advance to segment the foreground area, which bypasses the subject adjustment of parameters and can be applied to all types of organelles’ images (Fig. S7). Furthermore, the masks for nuclear membrane, nuclear reticulum, plasma membrane, filopodia, nucleus, cytosol, and ECS, are manually annotated on the super-resolution images.

Segmentation of intracellular compartments with DCNN networks

We train 10 DCNN networks to segment 15 intracellular compartments (Fig. 1, supplementary Movie 1). Namely, LD-net is responsible for recognizing lipid droplets, GOLGI-net for Golgi apparatus, MITO-net for mitochondria, PERO-net for peroxisome, EE-net for early endosome, LE-net for late endosome, LYSO-net for lysosome. ER-net is used to recognize the three structures of endoplasmic reticulum (ER), nuclear membrane, and nuclear reticulum, PM-net is used to recognize plasma membrane and filopodia, and VOLUME-net is to recognize other three structures of nucleus, cytosol, and extracellular space (ECS).

The super-resolution ratiometric images and ground truth masks are trained with the networks based on the attention U-Net architecture^{15, 16} (Fig. S6a). Golgi-net, LD-net, MITO-net, PERO-net, EE-net, LE-net, and LYSO-net are binary segmentation networks optimized with sigmoid cross-entropy loss, while ER-net, PM-net, and VOLUME-net are multi-class segmentation networks optimized with softmax cross-entropy loss. We segment each structure with a binary mask and the intensity images are generated by multiplying the Nile Red image with the binary mask (Fig. 2a, b). Compared with the prediction of multiple intensity levels, binary segmentation is simpler and more robust because only true/false decisions are made. Moreover,

the intensities of different colocalization probes varies across cells (Fig. S5a) when the fluorescent intensities of Nile Red are rather constant. Directly training with the intensity images will introduce errors. The output binary masks bring alternative advantages, which directly allow further quantitative analysis, such as organelles' number, volume, and contact frequency (Fig. S8). These information makes it possible to generate the digital anatomy and dynamics of live cells.

During training, we further perform the data augmentation on the fly to enlarge the training dataset and make the neuronal networks robust to different experimental conditions. The intensity images, ratio images, and corresponding binary masks are randomly rotated and scaled from 0.9x to 1.1x. The intensity images are randomly multiplied by 0.75x to 1.5x. Then the input images are resized and cropped into 2D or 3D patches to fit the network input. Since the abundance of different organelles also varies, we assign different class weights for different DCNN networks. The details of colocalization probes, training datasets, hyperparameters, and segmentation accuracies for every intracellular structure are provided in Supplementary Table 1 for the 3D dataset and in Supplementary Table 2 for the 2D dataset.

Evaluation of the segmentation performance

By comparing the masks between the ground truth and network output in Fig. 2a, b for the 3D datasets, the merged images and their quantitative analysis show the overall pixel accuracies of over 91.7% have been achieved for all the intracellular structures (Fig. 3a). We further evaluate the F1 score of Golgi apparatus (0.66), lipid droplets (0.72), mitochondria (0.63), peroxisomes (0.68), early endosomes (0.59), late endosomes (0.51), and lysosomes (0.67) predicted by the binary segmentation networks, as well as their recall and precision (Fig. 3b). For the multi-class segmentation networks of ER-net, PM-net, and VOLUME-net, the confusion matrixes show the portion of correctly predicted pixels is larger than 60% in most cases (Fig. 3d).

Since the DCNN networks predict the multi-color images of subcellular compartments from the ratiometric images, we further evaluate the performance of the "in-silico" multi-color images. The Pearson's correlation coefficient (PCC) and the Manders' overlapping coefficient (MOC)¹⁷ are metrics widely used in multi-color colocalization experiments, where MOC is more suitable to quantify colocalization of binary images. Comparing the predicted mask and the ground truth mask for each structure, we show that ER, lipid droplet, cytosol, nucleus, and ECS display strong colocalizations with the ground truth (MOC > 0.7), and all the other structures show good colocalizations (MOC > 0.5) (Fig. 3c). These results show that the predicted images from Nile Red fluorescence are well colocalized with the fluorescence images of the organelle-specific probes.

Compared with organelle segmentation on images of other modalities, our results show superior or comparable performance. For example, the predicted fluorescence images of Golgi apparatus on bright field images³ only reaches a low PCC value of less than 0.2. The Golgi mask segmented by our method reaches the MOC value of 0.67. According to previous studies¹⁷, the MOC and PCC have similar values when evaluating co-localization of binary images. Although predicted fluorescent images from the bright

field images can achieve PCC values of ~ 0.7 for mitochondria, the ultrafine structures such as the cristae cannot be resolved³. Compared with our results, organelle segmentation from EM dataset⁷ achieves slightly better performance on ER (F1 score = 0.80) and mitochondria (F1 score = 0.70), comparable performance on lysosomes (F1 score = 0.61), and worse performance on endosomes (F1 score = 0.39).

We further explore the critical role of ratio images in enhancing the segmentation accuracy for vesicle organelles with similar morphology. For example, the Ratio-net improves the F1 score than intensity-net in the dataset of early endosomes (0.59 vs 0.43, Fig. S9e-h) and Golgi apparatus (0.66 vs 0.54, Fig. S9i-l). For structures with distinct morphologies, such as mitochondria, the ratio image seems to add little value, as the F1 scores are similar for ratio-net and intensity-net (0.63 vs 0.62, Fig. S9a-d). Interestingly, the ratio-net and intensity-net also achieve similar F1 scores for lipid droplets (0.72 vs 0.69). It is possibly due to the distinct brightness of lipid droplets, as lipid droplets are the brightest organelles after Nile Red staining.

Evaluation of the segmentation performance transferred across instrument and cell types

We further evaluate the generalization power of the DCNN networks. Firstly, we test the performance of DCNN networks on different microscopes. The images of five different organelles are acquired on microscopes M1 and M2. Both microscopes are super-resolution spinning-disk microscopes, with slightly different excitation laser, spatial resolution and filter sets. When only trained with datasets from M1, the DCNN networks can segment the organelles but with lower accuracies. When trained with datasets from both M1 and M2, the prediction results show good colocalization ($\text{MOC} > 0.5$) with the ground truth masks (Fig. 3e). Secondly, we evaluate the segmentation performance on different cell types. Even only trained with U2-OS datasets, the DCNN networks can be used to predict COS7 and Hela datasets (Fig. S10). With transferred learning by only using a small number of COS7 and Hela images, the prediction accuracies can be significantly improved (Fig. 3f).

Moreover, we demonstrate that the proposed method can be applied to image the anatomic structures of living tissues. We show that Nile Red can homogeneously stain the membranous structures through the 3rd instar larvae testes, and in contrast, organic dyes, such as Mitotracker, can hardly penetrate cyst cells and diffuse homogeneously into spermatogonia (Fig. S1h). By conducting the colocalization experiments to recognize lipid droplets, lysosome (LysoView), and mitochondria (Mito-moxMaple3), and manually labeling the nucleus, cytosol, and extracellular space, as ground truth, we successfully obtained the six-color imaging of the intracellular organelles of complex multi-cellular sample of the entire testes system (Fig. 4f).

Cell anatomy imaging across mitotic stages.

According to the organization of intracellular structures, such as chromosomes, cytoskeleton, etc., the mitotic cell division experiences the four discrete phases: prophase, metaphase, anaphase, and

telophase^{18,19}. As the spatiotemporal rearrangements of membranous compartments are crucial to the propagation of cells, here we use the super-resolution in-silico multi-color imaging to study the 3D anatomy of live cells across mitotic phases. We demonstrated that the DCNN networks perform accuracy segmentations for most organelles during mitosis (Fig. 3f). As shown in Fig. 5a, Fig. 5b, and Fig. S11, the nuclei are clear in interphase, prophase, and telophase, become less obvious in anaphase, and almost disappear in metaphase. The nuclear reticulum can be resolved in interphase and prophase, as it promotes the disassembly of nuclear membrane¹⁹. In contrast, the plasma membrane becomes thicker and brighter since metaphase, which reflects their mechanical property changes (Fig. 5g). Interestingly, our results further quantitatively reveal the relatively small changes in the volume of the cells across different phases, ranging from 4000 μm^3 to 6000 μm^3 (Fig. 5h, Fig. S8a).

ER is the largest membranous organelle within the cell. The results presented in Fig. 5c and Fig. S11b show a denser ER structure around the nucleus with a mixture of tubule and sheet structures in interphase, while it has more tubule structures uniformly distributed in metaphase and anaphase. Figure 5d and Fig. S8c accordingly shows the similar trend for Golgi apparatus from being highly aggregated around the nucleus in interphase and telophase to the disassembly stage being segregated into smaller vesicles in prophase, metaphase and anaphase, and with a decreased total volume. The super-resolution in-silico images in Fig. 5e further reveal the elongated and stretched mitochondria in interphase and prophase compared with the fragmented structure in metaphase ready for equal inheritance²⁰. The total volume analysis of mitochondria shows the decrease from interphase to mitosis.

The 3D digital cell anatomy, as shown in Fig. 5 and supplementary Movie 2–6, further allows a statistical analysis of all the intracellular structures. For example, Fig. S11b plotted the spatial distributions of different compartmental pixels to the coverslip or nucleus, and Fig. S8 and Fig. S12 showed the particle volume, number, sphericity, and emission ratio of vesicular organelles, including lipid droplet, lysosome, early endosome, late endosome, and lysosome in different mitotic stages.

Fast imaging of the organelle interactome

Next, we apply the deep learning enhanced super resolution segmentation of membranous structures to study the intracellular organelles' interactomes, as the organelles coordinate with each other and undergo the spatiotemporal interactions to perform cellular functions. As the 3D volumetric imaging is insufficient in speed in capturing the dynamic interactions, we further use only 2D images within the first 1.6 μm to the coverslip to train the DCNN networks (Fig. S13). Our method offers a high throughput of 9 colors per exposure including 8 types of organelles and microtubules labeled by tubulin GFP, as the multiple detection channels are projected onto separate areas of a sCMOS camera (Fig. S2a). In comparison, the linear unmixing method based on spectral detection requires 26 images to solve 6 structures², resulting in the throughput of 0.23 color per exposure. The linear unmixing method based on excitation requires 6 or 8 images to observe 6 structures^{1,2}, resulting in the speed of 1.0/0.75 color per exposure. Therefore, in-

silico multicolor imaging, presented in our work, can increase the imaging speed by almost one order of magnitude.

With the two-color images of microtubule and Nile Red fluorescence images presented in Fig. 6a, the intracellular organelles can be further segmented from the Nile Red fluorescence images as shown in Fig. 6b-e. In the enlarged four-color image series of the microtubule, lipid droplet, mitochondria, and lysosome (Fig. 6c), the fission and fusion of mitochondria, the interaction between lipid droplet and microtubule, and the interaction among mitochondria, lysosome, and microtubule can be resolved. In Fig. 6d, the Golgi vesicles fuse and interact with the late endosome, and both types of organelles interact with the microtubule. In Fig. 6e, ER interacts with lipid droplets and peroxisomes. Supplementary Movie 7 shows the multiple interactions among 9 subcellular structures simultaneously imaged.

Discussion

Limitation of our technique

The prediction of DCNN networks always has errors. Here, to minimize the influence of errors, we choose to predict the binary masks instead of intensity images to avoid the generation of fake signals. We further train the DCNN networks with a variety of specimen and imaging data from different microscopes, from different cell lines, and on different cell cycles. Careful considerations of the error rates of the predicted results are still needed before biological conclusions can be meaningfully made. The MOC/F1 score varies across different cellular compartments, typically ranging from 0.5 to 0.7 for 3D datasets and from 0.3 to 0.6 for 2D datasets. For the low fidelity results, cross validation experiments are still needed. Besides, the networks should better be fine-tuned with additional training dataset under similar conditions in the experiments when using the pre-trained DCNN networks .

Potential application of our technique

We succeeded in resolving the digital anatomy of live cells by deep learning segmentation of the super-resolution data acquired on commercial microscopes. Only one dye, one excitation laser, and two-color images are found sufficient in resolving the 15 intracellular compartments and their dynamic interactions. Our method fundamentally overcomes the limitations of color crosstalks and the low joint labeling efficiency of multiple probes in the spectral domain. It significantly increases the imaging speed and throughput and reduces the photobleaching issues in live-cell imaging, which suggests the suitability of using the denoising power of deep learning^{21, 22} for long-term observation. It can also accelerate the multi-color imaging of a large field of view (Fig. S14).

While lipid dyes have gained considerable interest in subcellular imaging^{14, 23-26}, the specificity of lipid dyes becomes a major challenge, due to the large variety of lipid molecules and membrane structures existing within the cell. Our method of deep learning segmentation gears the disadvantage of unspecific staining into an advantage towards the multiplexed imaging of complex intracellular and multi-cellular structures. Moreover, our method further benefits from the other merits of lipid dyes, including the efficient

staining, the easy use in a push-pull manner, the exchangeable nature suitable for long-term observation, and the responsive feature for functional observation of physicochemical properties of subcellular membranes. Our work further suggests that a variety of recently developed lipid dyes^{24, 26–28} should be compatible with our method. Taking these together, the features with high brightness and biocompatibility, and a heterogeneous and responsive emission spectrum across different organelles, are highly favorable to improve the segmentation accuracies in resolving the organelles with similar morphology.

Deep learning segmentation of organelles can also be achievable in a broad range of super resolution microscopes, as the lipid dyes have been demonstrated suitable in Structured Illumination Microscopy (SIM)¹⁴, STimulated Emission Depletion (STED)²³, and Single Molecule Localization Microscopy (SMLM)^{25, 28, 29}. To reduce phototoxicity, adoption of our method to long-term volumetric imaging techniques, such as lattice light sheet³⁰, may lead to the tracking of organelle dynamics across mitotic stages of cells in their native environment.

Our method is also orthogonal to other existing approaches, such as multi-color labeling and spectrum unmixing, so that the level of complexity can be further increased by combining other fluorescent probes to study the subcellular interactions among organelles, cytoskeleton, and single molecular proteins. Besides the fluorescence intensity and wavelength, the fluorescence polarization¹⁴ and lifetime^{31, 32} may serve as alternative “optical fingerprints” for the DCNN networks to further increase the segmentation accuracy and imaging throughput.

Methods

Sample Preparation. Human osteosarcoma U2-OS cell lines (HTB-96, ATCC, USA) were cultured in Dulbecco's Modified Eagle's medium (DMEM, GIBCO, USA) containing 10% heat-inactivated fetal bovine serum (FBS, GIBCO, USA) and 100 U/ml penicillin and 100 µg/ml streptomycin solution (PS, GIBCO, USA) at 37°C in an incubator with 95% humidity and 5% CO₂. For the living cell imaging, the cells were plated at the desired density on the µ-Slide 8 Well (80827, ibidi, USA) and 1 µg/ml Nile Red (N1142, Invitrogen, USA) was added into the culture medium 1 h before imaging and was present during imaging.

For colocalization experiments, the cells were transfected 16 h before imaging with the plasmids of early endosome-GFP (Rab5a, C10586, BacMam 2.0, CellLight, USA), late endosome-GFP (Rab7a, C10588), ER-GFP (ER signal sequence of calreticulin and KDEL, C10590), Golgi-GFP (Golgi-resident enzyme N-acetylgalactosaminyltransferase 2, C10592), Mitochondria-GFP (leader sequence of E1 alpha pyruvate dehydrogenase, C10600) and incubated overnight. 1 µg/ml Nile Red was added into the culture medium 1 h before imaging. The colocalization of lysosome is obtained by Lysoview™ 488 (70067-T, Biotium) 30 min before imaging and without washing during imaging. For mitosis experiments, the nuclei were stained by NucSpot® Live 488 Nuclear Stain (40081-T, Biotium) 10 min before imaging. The verapamil was added to the staining solution to improve probe retention in the live cell. The microtubule was transfected by Lipofectamine™ 3000 (L3000-001, Invitrogen™) with MAP4-sfGFP 36 h before imaging.

For live-cell imaging of *Drosophila* testes, the testes from 3rd instar larvae were dissected in PBS and then placed into a confocal imaging dish containing the culture medium and diluted Nile Red. After 15 minutes, the testes stained with Nile Red can be imaged under fluorescent microscopy. For colocalization imaging of lysosomes and lipid droplets, we used wild-type *Drosophila* (W118). For colocalization imaging of mitochondria, we used transgenic *Drosophila* (UASp-mito-Moxmaple3) crossed by the GAL4/UAS system.

System setup. Two microscopes which are slightly different are used in our experiments. Microscope 1: an inverted fluorescence microscope (TI-E, Nikon) is equipped with a TIRF objective (CFI Apochromat TIRF $\times 100$ oil, NA 1.49, Nikon) and a spinning disk confocal system (CSU-W1, Yokogawa). The super-resolution imaging module (Live SR, Gataca) could double the imaging resolution. A 488 nm laser (iLAS 3) is used for excitation. Emission fluorescence was acquired by sCMOS camera (Prime 95B, Photometrics) after different emission filters (green: S525/50m; yellow: S617/73m; red: S685/40m; Chroma) in the spinning disk system. Microscope 2: an inverted fluorescence microscope (TI-2, Nikon) equipped with a TIRF objective (CFI Apochromat TIRF $\times 60$ oil, NA 1.49, Nikon) and a spinning disk confocal system (X-Light V3, CrestOptics). A 470 nm laser (LDI, 89 North) is used for excitation and the super-resolution imaging module (Live SR, Gataca) could double the imaging resolution. Emission fluorescence passes through a multi-splitter module (CAIRN, MultiSplit V2) with three dichroic mirrors (ZT488rdc/ZT561rdc/ZT640rdc; Chroma) and different emission filter (blue: ET480/40m; green: S525/50m; yellow: S617/73m; red: S685/40m; Chroma). Then illuminated on the different parts of the camera (Prime BSI, Photometrics). Most of the datasets were acquired from Microscope 2. Half the 3D datasets of Golgi apparatus, lipid droplets, mitochondria, and ER are acquired from Microscope 1. The acquisition process is performed on the Metamorph software. The resolution of the system is analyzed by PSFj software (Fig. S2) and Fourier Ring Correlation (Fig. S3).

Training data preparation. Multi-channel images are registered with chromatic aberration correction algorithm³³. The emission ratio was calculated by dividing the red channel by the yellow channel. The fluorescence image is the average image between the red channel and yellow channel. For the 3D dataset, the whole imaging FoV is $1200 \times 1200 \times N_z$ px³ (N_z ranges from 12 to 30), which is resized into $802 \times 802 \times 24$ px³ and is further cropped into random $256 \times 256 \times 24$ px³ patches. For the 2D dataset, the images are resized and cropped into random 256×256 px² patches.

For the structures of the nuclear membrane, nuclear reticulum, plasma membrane, filopodia, nucleus, cytosol, and ECS, the ground truth masks are manually annotated. For other structures, the green-channel colocalization images are acquired. In green-channels images, Nile Red shows strong fluorescence in LD and weak fluorescence in other organelles. To exclude the weak Nile Red fluorescence of other organelles, the green-channel images are subtracted by a portion of the yellow-channel images ($ChG-r \times ChY$, $r = 0.05$). To exclude the strong Nile Red fluorescence in LD, the pixels within the LD mask, which is predicted by the pre-trained LD network, are set to zero intensity. The ground truth masks are segmented from the colocalization image by a pre-trained FG-net (Fig. S6b), which is trained with 40 manually annotated colocalization images of various organelles.

Model architecture and training. Our DCNN networks are based on the attention U-Net architecture and are implemented in Python using the PyTorch package. The FG-net takes 2D intensity images as input (1@256*256) (Fig. S6b), and the subcellular segmentation networks take 3D intensity and ratio images as input (2@256*256*24) (Fig. S6a). Both binary segmentation networks and multi-class segmentation networks are optimized by Adam optimizer with cross-entropy loss functions. The starting learning rate of 0.0001, which is reduced on the plateau with the factor of 0.3 and patience of 5. The class weights of different organelles are adjusted according to the abundance of corresponding organelles. For example, a cell consists of more pixels labeled as mitochondria than as peroxisome, so that the weight of mitochondria is 2 and the weight of peroxisome is 5 (Supplementary Table 1). The number of training patches varies from 160 ~ 204 for different structures, which are trained on the RTX 2080 Ti GPU within 12 h. Compared with the 3D dataset, the class weights and data configuration in the 2D dataset is slightly different (Supplementary Table 2).

Image analysis and display. The evaluation of predicted masks is performed by custom-written Matlab. The predicted masks for each patch are merged and resized back, and the metrics between the ground truth and prediction are calculated as follows:

$$\text{Pixel accuracy} = \frac{N_{tp} + N_{tn}}{N_{tp} + N_{fp} + N_{tn} + N_{fn}}$$

$$\text{Recall} = \frac{N_{tp}}{N_{tp} + N_{fn}}$$

$$\text{Precision} = \frac{N_{tp}}{N_{tp} + N_{fp}}$$

$$\text{F1 score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

$$\text{MOC} = \frac{\sum xy}{\sqrt{\sum x^2 + \sum y^2}}$$

where N_{tp} , N_{fp} , N_{tn} , N_{fn} are the number of pixels for true-positive, false-positive, true-negative, and false-negative; x , y are pixel values of the predicted mask and the ground truth mask. For the confusion matrixes, the value in the i th row, j th column (N_{ij}) is the number of pixels with label i in the ground truth image and label j in the predicted image.

The 3D intensity images are generated in Imaris software (Bitplane) by volumetric intensity rendering at MIP mode and the 3D volume of binary masks is generated by surface rendering. The results of the statistical (Fig. S8, 11–12) are generated by Imaris, and drawn in Prism and Origin. The volume, number, and sphericity are calculated by the surface tool with predicted binary masks. Objects smaller than 10 voxels were excluded from the statistics. The average value of the ratio in each surface object is calculated to draw the 2D scatter plot (Fig. S12). The spot tool in Imaris was used to find the center of the object to calculate the distance from the nucleus and the coverslip (Fig. S11).

Code and Data availability

All data that support the findings of this study will be publicly available with the publication of the paper. We are glad to provide the data for review purpose upon request.

Code availability

The code used in this paper is available on <https://github.com/KarlZhanghao/LiCAI>.

Declarations

Acknowledgements

This work was supported by the National Natural Science Foundation of China (62005116, 62025501, 31971376, 92150301), and Shenzhen Science and Technology Program (RCBS20200714114817138, KQTD20170810110913065). We thank National Center for Protein Sciences at Peking University in Beijing, China, for assistance with the Gataca Live SR super-resolution imaging. We thank Dr. Siying Qin of Peking University for assistance with 3D image analysis using Imaris. We thank Prof. Kuangshi Chen of Peking University for his generosity in providing cells.

References

1. Chen, K., Yan, R., Xiang, L.M. & Xu, K. Excitation spectral microscopy for highly multiplexed fluorescence imaging and quantitative biosensing. *Light-Science & Applications* **10**, 97 (2021).
2. Valm, A.M. et al. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017).
3. 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**, 917–920 (2018).
4. Christiansen, E.M. et al. In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images. *Cell* **173**, 792–803 (2018).
5. Manifold, B., Men, S., Hu, R. & Fu, D. A versatile deep learning architecture for classification and label-free prediction of hyperspectral images. *Nature Machine Intelligence* **3**, 306–315 (2021).
6. Cheng, S.Y. et al. Single-cell cytometry via multiplexed fluorescence prediction by label-free reflectance microscopy. *Science Advances* **7**, eabe0431 (2021).
7. Heinrich, L. et al. Whole-cell organelle segmentation in volume electron microscopy. *Nature* **599**, 141–146 (2021).
8. Nixon-Abell, J. et al. Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. *Science* **354**, aaf3928 (2016).

9. Shroff, H., Galbraith, C.G., Galbraith, J.A. & Betzig, E. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nature Methods* **5**, 417–423 (2008).
10. Huang, X. et al. Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy. *Nature Biotechnology* **36**, 451–459 (2018).
11. Yang, X.S. et al. Mitochondrial dynamics quantitatively revealed by STED nanoscopy with an enhanced squaraine variant probe. *Nature Communications* **11**, 3699 (2020).
12. Guo, Y. et al. Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution on Millisecond Timescales. *Cell* **175**, 1430–1442 e1417 (2018).
13. Azuma, T. & Kei, T. Super-resolution spinning-disk confocal microscopy using optical photon reassignment. *Opt Express* **23**, 15003–15011 (2015).
14. Zhanghao, K. et al. High-dimensional super-resolution imaging reveals heterogeneity and dynamics of subcellular lipid membranes. *Nature Communications* **11**, 5890 (2020).
15. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 234–241 (2015).
16. Oktay, O. et al. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv* (2018).
17. Manders, E.M.M., Verbeek, F.J. & Aten, J.A. Measurement of co-localization of objects in dual-colour confocal images. *Journal of Microscopy* **169**, 375–382 (1993).
18. Jongsma, M.L., Berlin, I. & Neefjes, J. On the move: organelle dynamics during mitosis. *Trends in Cell Biology* **25**, 112–124 (2015).
19. Carlton, J.G., Jones, H. & Eggert, U.S. Membrane and organelle dynamics during cell division. *Nat Rev Mol Cell Biol* **21**, 151–166 (2020).
20. Moore, A.S. et al. Actin cables and comet tails organize mitochondrial networks in mitosis. *Nature* **591**, 659–664 (2021).
21. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature Methods* **15**, 1090–1097 (2018).
22. Li, X.Y. et al. Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising. *Nature Methods* **18**, 1395–1400 (2021).
23. Spahn, C., Grimm, J.B., Lavis, L.D., Lampe, M. & Heilemann, M. Whole-Cell, 3D, and Multicolor STED Imaging with Exchangeable Fluorophores. *Nano Letters* **19**, 500–505 (2019).
24. Collot, M. et al. Probing Polarity and Heterogeneity of Lipid Droplets in Live Cells Using a Push-Pull Fluorophore. *Analytical Chemistry* **91**, 1928–1935 (2019).
25. Moon, S. et al. Spectrally Resolved, Functional Super-Resolution Microscopy Reveals Nanoscale Compositional Heterogeneity in Live-Cell Membranes. *Journal of the American Chemical Society* **139**, 10944–10947 (2017).
26. Darwich, Z., Klymchenko, A.S., Dujardin, D. & Mély, Y. Imaging lipid order changes in endosome membranes of live cells by using a Nile Red-based membrane probe. *RSC Advances* **4**, 8481–8488

(2014).

27. Collot, M. et al. MemBright: A Family of Fluorescent Membrane Probes for Advanced Cellular Imaging and Neuroscience. *Cell Chemical Biology* **26**, 600–614 e607 (2019).
28. Kucharak, O.A. et al. Switchable Nile Red-Based Probe for Cholesterol and Lipid Order at the Outer Leaflet of Biomembranes. *Journal of the American Chemical Society* **132**, 4907–4916 (2010).
29. Sharonov, A. & Hochstrasser, R.M. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences* **103**, 18911–18916 (2006).
30. Chen, B.C. et al. Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014).
31. Lu, Y. et al. Tunable lifetime multiplexing using luminescent nanocrystals. *Nature Photonics* **8**, 32–36 (2013).
32. Castello, M. et al. A robust and versatile platform for image scanning microscopy enabling super-resolution FLIM. *Nature Methods* **16**, 175–178 (2019).
33. Au - Matsuda, A., Au - Koujin, T., Au - Schermelleh, L., Au - Haraguchi, T. & Au - Hiraoka, Y. High-Accuracy Correction of 3D Chromatic Shifts in the Age of Super-Resolution Biological Imaging Using Chromagnon. *Journal of Visualized Experiments*, e60800 (2020).

Supplementary Movies

Supplementary Movies 1-7 are not available with this version.

Figures

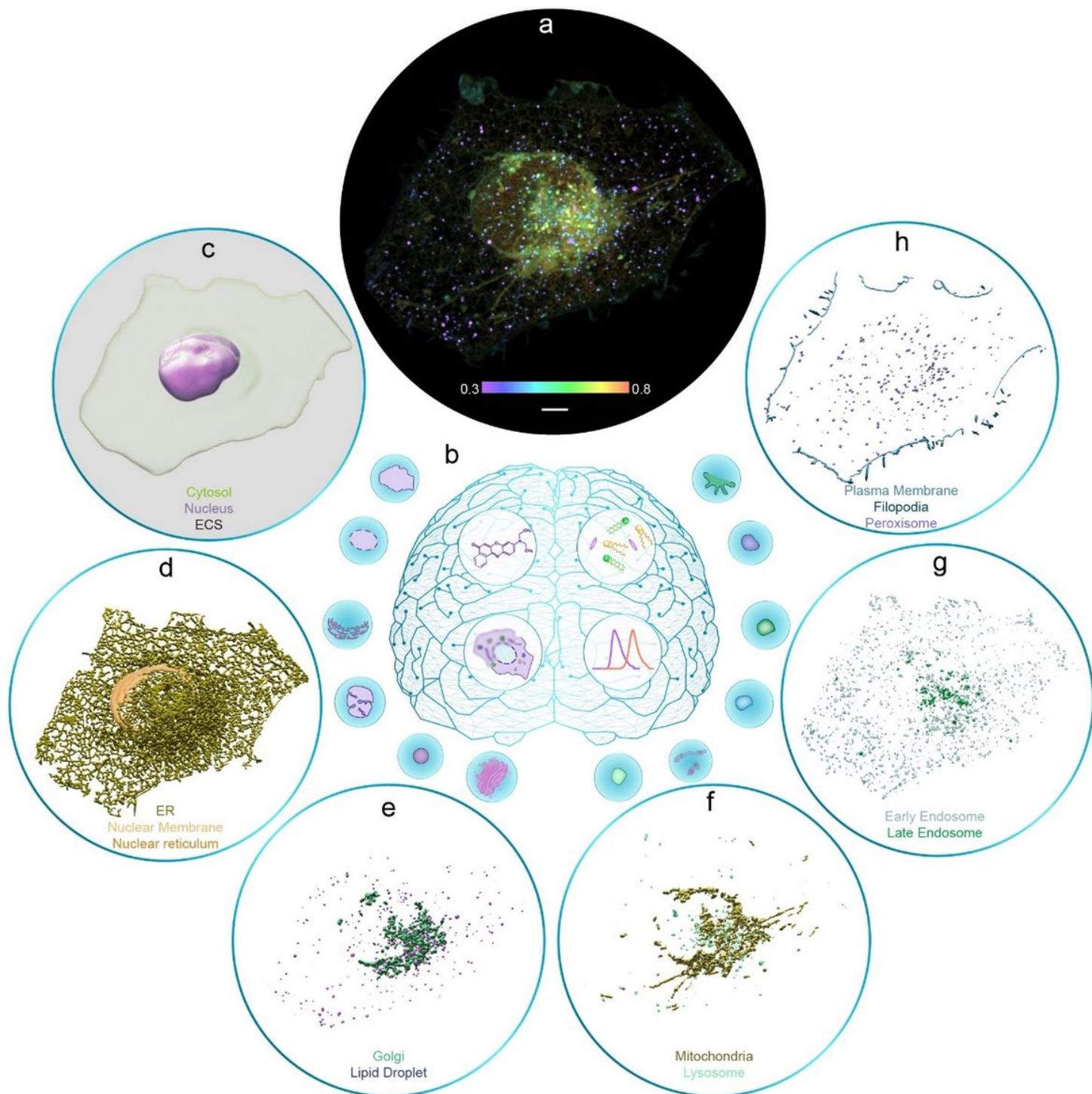


Figure 1

Super-resolution imaging of live cell anatomy by the universal lipid staining and deep learning. (a) Super-resolution image of a U2-OS cell with its membranous structures stained by Nile Red. The pseudo color represents the spectral ratio between the yellow channel (Em: 580-653 nm) and the red channel (Em: 665-705 nm) under single laser excitation at 488 nm. **(b)** From the super-resolution intensity image and the spectral ratio image, the deep convolutional neural networks (DCNN) predict binary masks of 15 subcellular structures. **(c)** The masks of cytosol, nucleus, and extracellular space are segmented by

VOLUME-net. **(d)** The masks of ER, nuclear reticulum, and nuclear membrane are segmented by ER-net. **(e)** The Golgi mask is segmented by GOLGI-net and the lipid droplet mask is segmented by LD-net. **(f)** Mitochondria mask segmented by MITO-net and lysosome mask segmented by LYSO-net. **(g)** Early endosome mask segmented by EE-net and late endosome mask segmented by LE-net. **(h)** Peroxisome mask segmented by PERO-net and the masks of the plasma membrane and filopodia segmented by PM-net. Scale bar: 5 μ m.

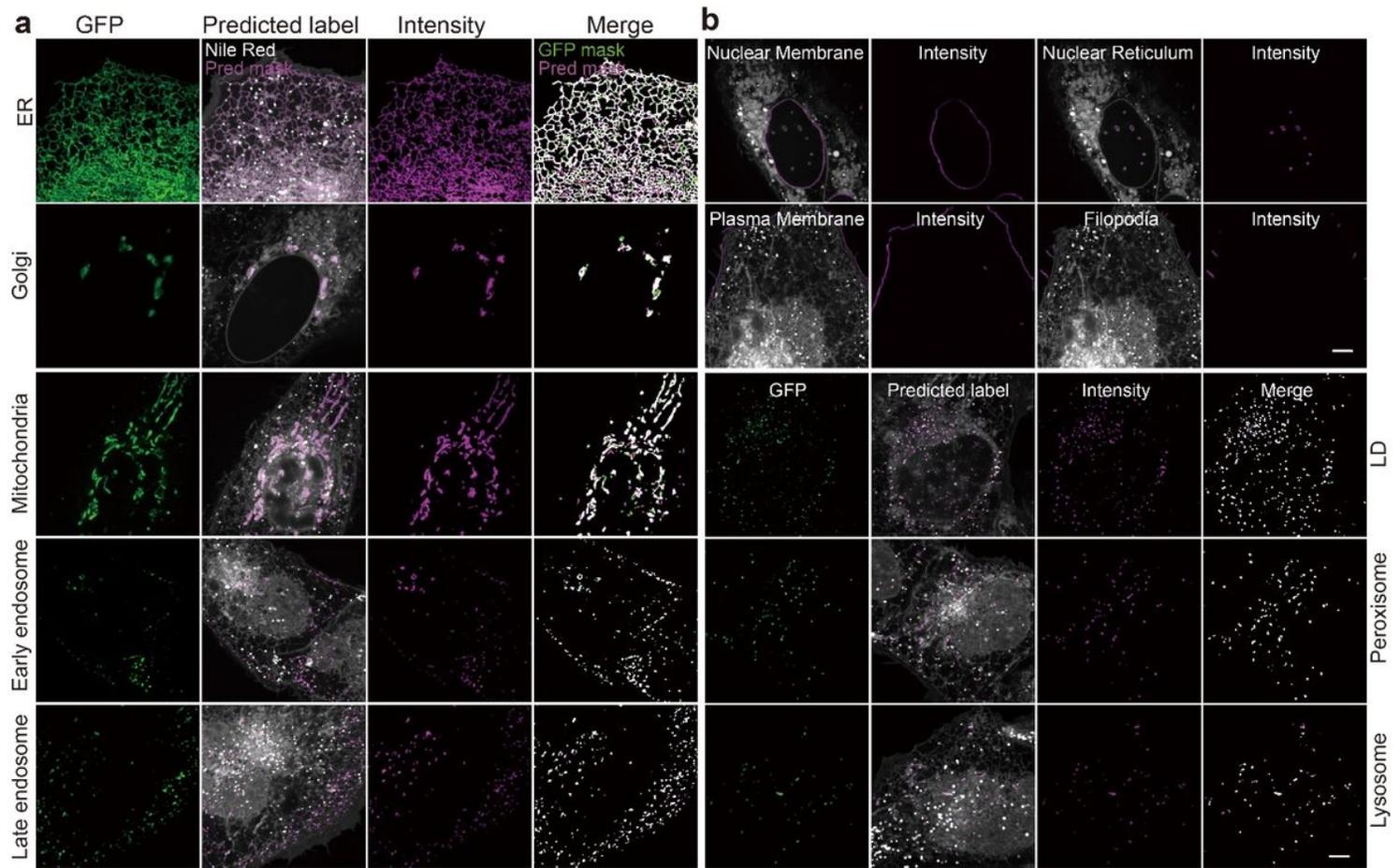


Figure 2

Accurate in-silico prediction of the 15 intracellular structures. **(a)** The first column shows the GFP colocalization images; the second column shows the Nile Red fluorescence images with magenta masks predicted by the networks; the third column shows the intensity images by multiplying the fluorescence image and the predicted masks; the fourth column shows the merged images of predicted masks (in magenta) and colocalization masks (in green, FG-net prediction of colocalization intensity images). In the merged images, white or black pixels indicate true prediction; magenta pixels indicate wrong prediction (false positive) and green pixels indicate missing prediction (false negative). **(b)** The predicted masks of the nuclear membrane, nuclear reticulum, plasma membrane, and filopodia on the fluorescence images and the masked intensity images. Scale bar: 5 μ m.

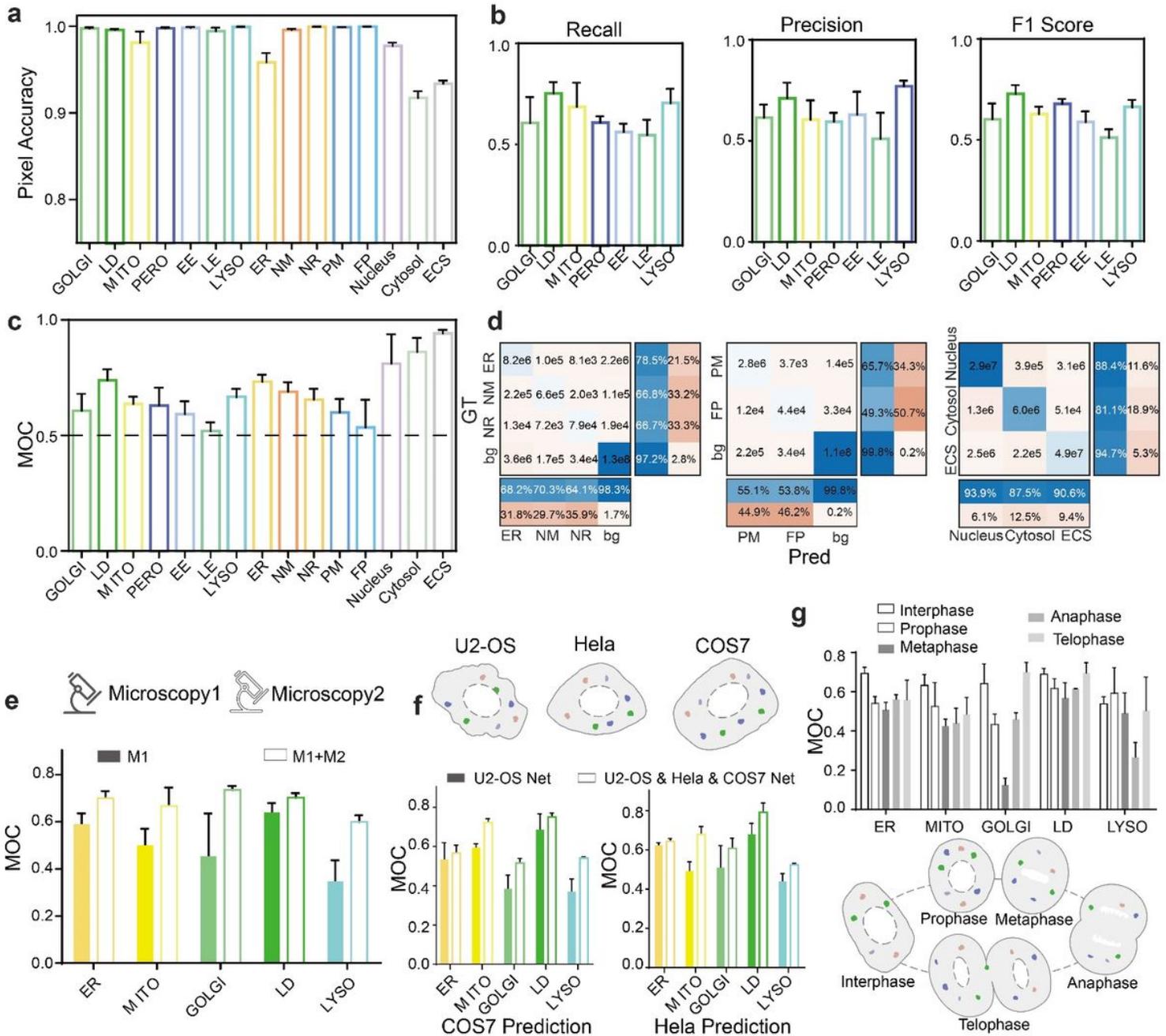


Figure 3

Segmentation accuracies of different DCNN networks. (a) Pixel accuracies of 15 structures. **(b)** Recall, precision, and f1 score of binary segmentation networks. **(c)** The Manders' overlapping coefficient (MOC) of 15 structures between predicted masks and ground-truth masks. **(d)** Confusion matrixes of multi-classification networks including ER-net, PM-net, and VOLUME-net. **(e)** The MOC of datasets from different microscopy. The organelle images are acquired from microscope M2 and predicted by two DCNNs. One DCNN only used the training datasets acquired from microscope M1 (solid bars), and the other DCNN used the training datasets from both M1 and M2 (hollow bars). The results demonstrate the generalization of the DCNN networks on different microscopes. **(f)** The MOC from different cell types. Even only trained with U2-OS datasets, the DCNN networks can predict COS7 or HeLa datasets with

moderate colocalization. With transferred learning with a small number of COS7 and Hela images, the prediction accuracies can be significantly improved. **(g)** Prediction accuracies of the DCNN network on different mitosis stages.

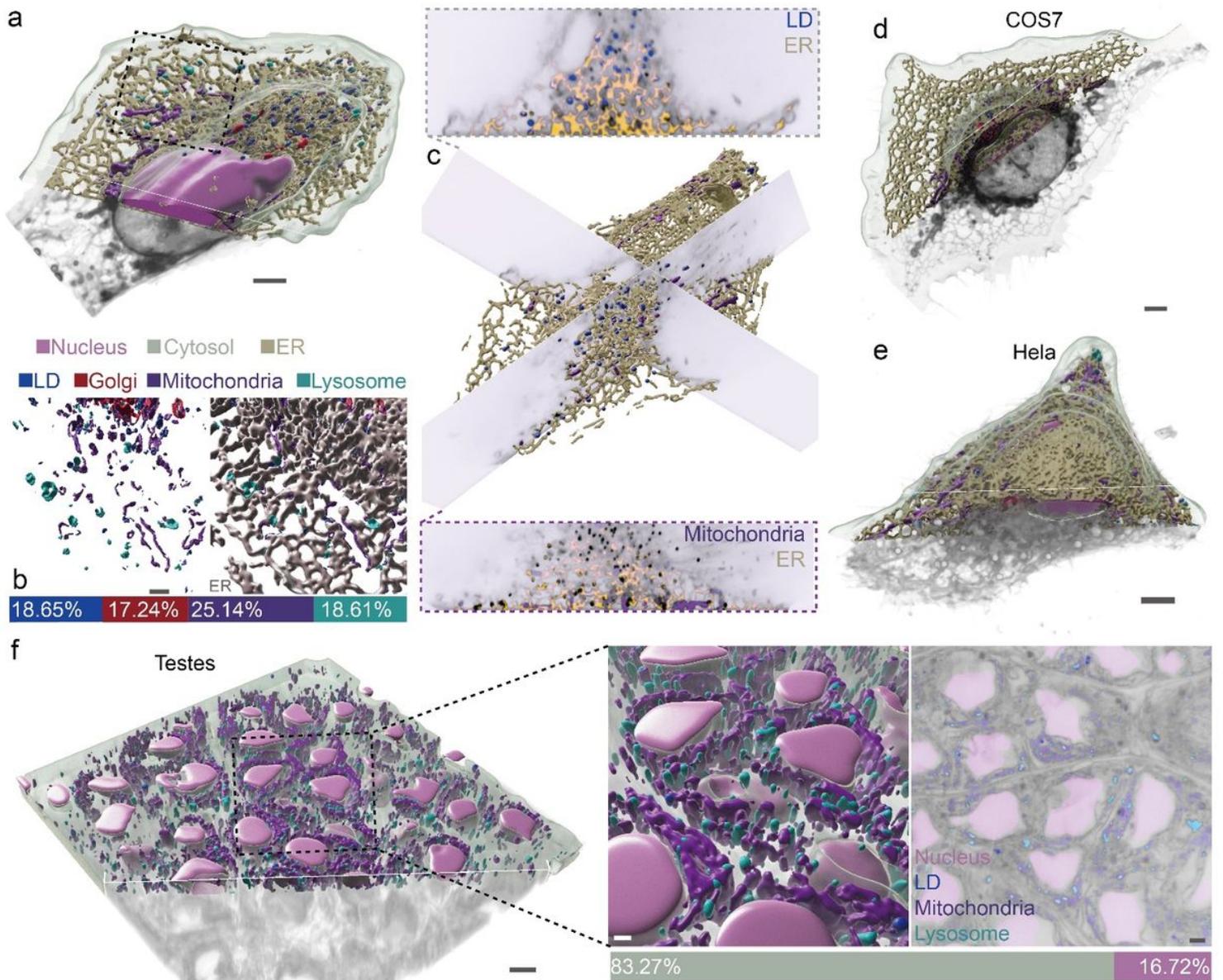


Figure 4

Organelle segmentation on different biology samples by the universal lipid staining and deep learning. (a) Segmentation results of U2-OS cells. **(b)** Zoomed-in view in the black box in (a), which shows interactions of lipid droplet, Golgi, Mitochondria, lysosome with ER. The bar shows the portion of interaction area relative to the total volume of the corresponding organelle. **(c)** Two different cross-sectional images and surface rendering images show the segmentation results overlaid with intensity image. **(d-e)** 3D rendering of organelle segmentations for COS7 and Hela cell types. **(f)** The anatomic structures segmented from living tissue images of 3rd instar larvae testes in *drosophila* by the universal lipid staining and deep learning. Scale bar: 5 μ m.

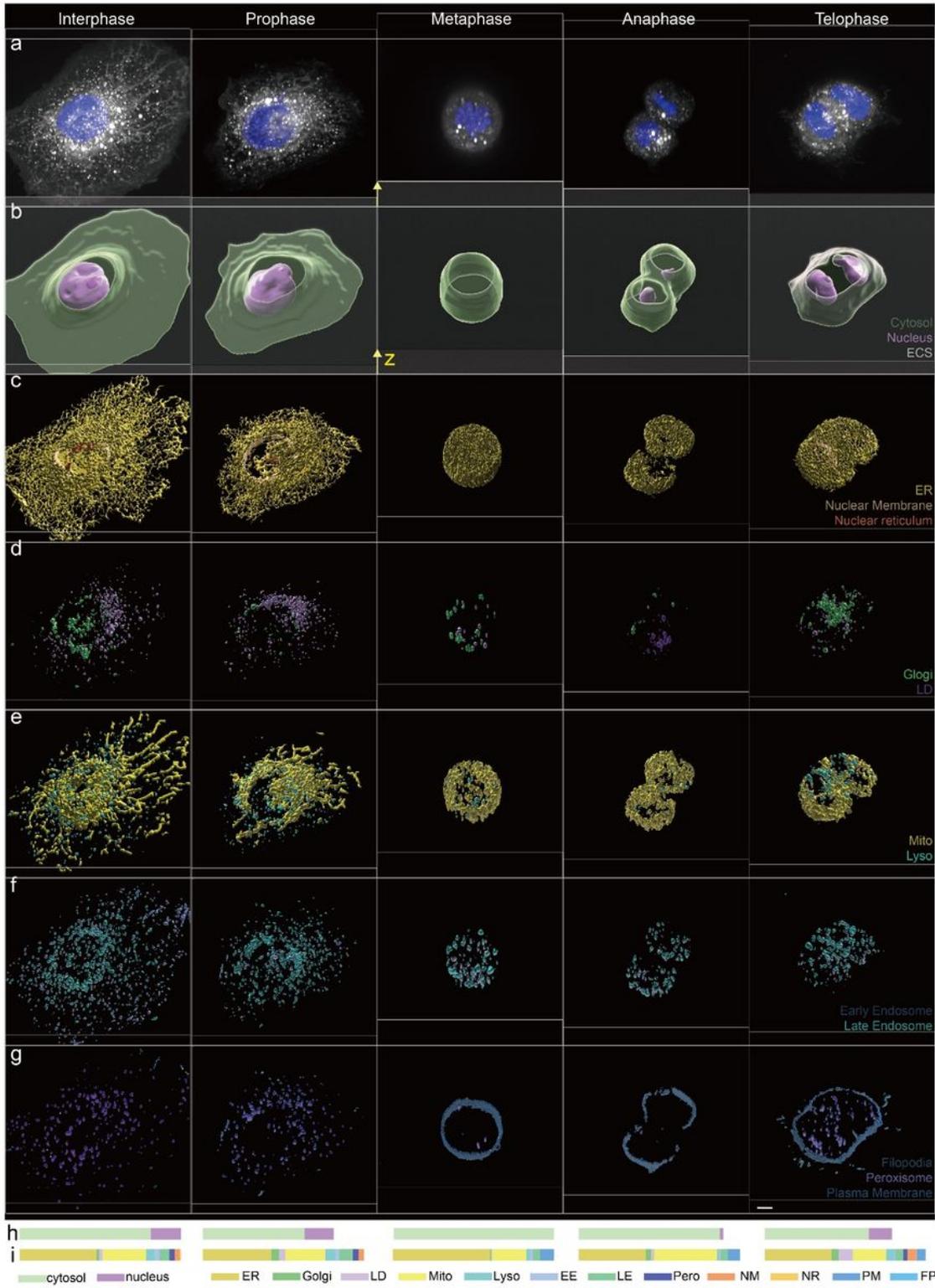


Figure 5

Dynamic imaging of cell anatomy across all the mitosis stages. (a) Super-resolution image of a U2-OS cell with its membranous structures stained by Nile Red and the nuclei structure stained by NucSpot® live dye. (b-g) The arrangement of intracellular structures including cytosol, nucleus, ECS, ER, nuclear membrane, nuclear reticulum, Golgi, lipid droplet, mitochondria, lysosome, early endosome, late endosome, plasma membrane, filopodia, and peroxisome at interphase, prophase, metaphase, anaphase,

and telophase. The yellow arrow in (b) points from the bottom of the coverslip to the top of the cell. **(h)** The total volume analysis of cytosol and nucleus at different mitosis stages. **(i)** The volume proportion of ER, Golgi, lipid droplet, mitochondria, lysosome, early endosome, late endosome, peroxisome, nuclear membrane, nuclear reticulum, plasma membrane, and filopodia at different mitosis stages. The total volume calculates the foreground area of Nile Red fluorescence images. See Fig. S8 for detailed volume statistics of different intracellular structures. Scale bar: 5 μ m.

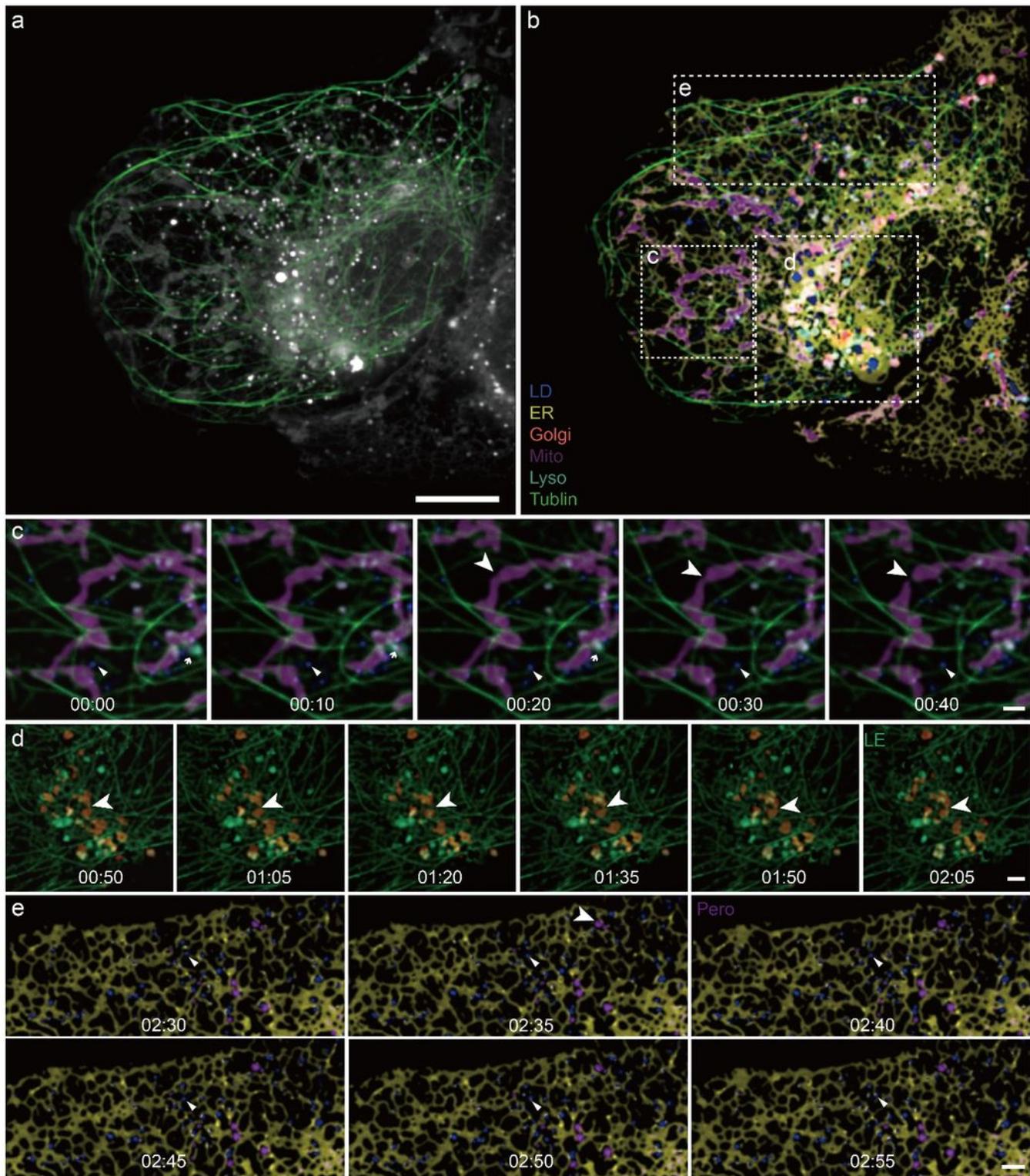


Figure 6

Fast, live-cell imaging of the dynamic interactions among multiple organelles. **(a)** The super-resolution images of microtubule (green) and membranous structures (gray). **(b)** Different types of organelles are segmented and the masks are applied to the intensity image, resulting in the multi-color image of lipid droplet, ER, Golgi apparatus, mitochondria, and lysosome. **(c)** Zoom-in time-lapse images of the interactions among microtubule, mitochondria, lipid droplet, and lysosome. **(d)** Zoom-in time-lapse images of the interactions among microtubule, Golgi apparatus, and late endosome. **(e)** Zoom-in time-lapse images of the interactions among ER, lipid droplet, and peroxisome. Scale bar: (a, b) 10 μm , (c, e) 5 μm , (d) 2 μm .

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [NatureArticleSIv1.8.docx](#)