Article

# Fast segmentation and multiplexing imaging of organelles in live cells

Karl Zhanghao [1,2,9], Meiqi Li [1,3,9] ✉, Xingye Chen[4], Wenhui Liu[4], Tianling Li[5], Yiming Wang[3], Fei Su[6], Zihan Wu[7], Chunyan Shan [3], Jiamin Wu [4], Yan Zhang [3], Jingyan Fu [5], Peng Xi [1,7,8] ✉ & Dayong Jin [2,6,7] ✉

Studying organelles' interactome at system level requires simultaneous observation of subcellular compartments and tracking their dynamics. Conventional multicolor approaches rely on specific fluorescence labeling, where the number of resolvable colors is far less than the types of organelles. Here, we use a lipid-specific dye to stain all the membrane-associated organelles and spinning-disk microscopes with an extended resolution of ~143 nm for high spatiotemporal acquisition. Due to the chromatic polarity sensitivity, high-resolution ratiometric images well reflect the heterogeneity of organelles. With deep convolutional neuronal networks, we successfully segmented up to 15 subcellular structures using one laser excitation. We further show that transfer learning can predict both 3D and 2D datasets from different microscopes, different cell types, and even complex systems of living tissues. We succeeded in resolving the 3D anatomic structure of live cells at different mitotic phases and tracking the fast dynamic interactions among six intracellular compartments with high robustness.

A cell has multiple intracellular compartments that interact with each other to form a systematic and self-efficient community. Organelles' interactome has become an emerging field of research[1,2]. Although up to 35 organelles throughout a whole cell have been annotated in volume electron microscopy at nanometer resolution[3], it is limited to fixed cells. Optical microscopy can image organelles in live cells, but is challenged by the small sizes, fast dynamics, and diverse species of the organelle types: the sizes of the ultrastructure of organelles are often less than the diffraction resolution limit of optical microscopy, the dynamic interactions of organelles within the crowded space pose motion artifact during the imaging process, and more than a dozen organelle types playing crucial roles in the function of a live cell, which must be monitored simultaneously.

Fluorescence microscopy has been widely used, as the fluorophores can be specifically labeled/expressed on the targeted subcellular proteins. But limited by the wide excitation and emission spectra of the fluorescence, it often requires different excitation and detection spectrum. Multiple excitation/emission steps will proportionally reduce the imaging speed. To address this problem, simultaneous excitation of multiple lasers and multi-spectral detection have been employed to observe up to six types of organelles and their interactions[1,4]. Further increase in color is intrinsically limited by the issue of dyes' crosstalk in the spectrum domain. The phototoxicity brought by the multiple laser excitation becomes another limiting factor for live cell imaging.

Moreover, multicolor labeling of sophisticated intracellular components by multiple fluorophores can be tedious, non-specific, and

[1]Department of Biomedical Engineering, College of Future Technology, Peking University, Beijing, China. [2]Zhejiang Provincial Engineering Research Center for Organelles Diagnostics and Therapy, Eastern Institute of Technology, Ningbo, Zhejiang, China. [3]School of Life Sciences, Peking University, Beijing, China. [4]Department of Automation, Tsinghua University, Beijing, China. [5]State Key Laboratory of Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing, China. [6]Institute for Biomedical Materials & Devices (IBMD), University of Technology Sydney, Sydney, NSW, Australia. [7]UTS-SUStech Joint Research Centre for Biomedical Materials & Devices, College of Engineering, Southern University of Science and Technology, Shenzhen, Guangdong, China. [8]National Biomedical Imaging Center, Peking University, Beijing, China. [9]These authors contributed equally: Karl Zhanghao, Meiqi Li. ✉e-mail: limeiqi@pku.edu.cn; xipeng@pku.edu.cn; dayong.jin@uts.edu.au

inefficient. Many organelles are surrounded by lipid membranes that share similar surface and chemical properties, and therefore, it is very difficult for organic dyes to specifically label and differentiate each organelle type. Though fluorescent proteins can achieve better specificity to label some types of organelles, the plasmids transfection often suffers from no expression or overexpression (Fig. S1a–f). The joint labeling strategy may also fail, as its efficiency will drop exponentially with the increase of multiple probes (Fig. S1g). The above-mentioned situations become even more challenging for both the diffusion of organic dyes and the expression of fluorescent proteins in living tissues (Fig. S1h, i).

It becomes particularly challenging for super-resolution (SR) fluorescence microscopy techniques to perform multiplexing imaging of organelles in live cells[5–8]. To date, only up to three colors have been demonstrated[9,10]. This is due to the fact that the improvement of spatial resolution is often at the scarification of the temporal resolution[11], and a balance between spatial resolution, sensitivity, phototoxicity, and imaging speed becomes necessary for live cell organelle interactome.

With the emergence of artificial intelligence, deep learning has been introduced to the in-silico prediction of multi-color images from transmitted bright field images[12–15], 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 existing training images suffer from low resolution and low contrast, the technique is yet to be satisfactory in predicting the subcellular structures. Compared to SR fluorescence imaging, label-free imaging is deficient in providing ample physical information, thereby impeding the study of multiple organelles' interactions. 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[12]. Although the predicted images of mitochondria (MITO) 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.

To explore the unknown subcellular world, the "predicted" extended-resolution images by DCNN should be based on the meaningful physical process. In a recent work[3], the whole-cell segmentation of up to 35 organelles has been achieved on volume electron microscopy, but it is limited in fixed cells. SR microscopy, by taking advantage of high resolution and contrast, can directly image the subcellular structures and dynamics[5–8]. As the data are obtained from the clear physical process, they can be directly used to train DCNN and to obtain the unknown behaviors of multiple organelles in live cells. Moreover, although long-term imaging in live cells has been demonstrated[7,8,16], only two or three colors can be achieved. Besides the problem of spectrum crosstalk, long-term imaging 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, reconstruct images, and increase spatial resolution, in-silico multicolor imaging of multiple organelles and subcellular compartments in SR microscopes remains unexplored.

Here, we report an approach to tackle the above listed challenges in multiplexing imaging of the intracellular organelles and their interactions. We abandon the conventional "one-to-one" fluorescence specific labeling strategy, and adapt a "one-to-many" strategy for organelles, i.e., to stain multiple intracellular compartments with one lipid dye that can stain the membranous structures with nearly 100% efficiency. We further use a spinning disk confocal microscope with an

extended resolution of ~143 nm[17] to obtain a series of images at high temporal resolution (Fig. S2). Moreover, as the emission spectrum of the dye responds to the lipid polarity of membranes[18,19], we further apply dual-color imaging channels at 617 and 685 nm bands to obtain a ratiometric measurement, so as to discriminate organelles with similar shapes and sizes. Our previous study has shown that Nile Red successfully labels at least ten subcellular compartments. The emission spectrum of Nile Red exhibits a red shift in polar environments and a blue shift in nonpolar environments, which can be linked to the composition of lipid membranes[20]. Here, we demonstrate such a simple cell staining protocol 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, speed, reproducibility, stability, and throughput, opening a method for multiplexing imaging inside the living cell and tissues (Fig. 1).

## Results

### The robustness of nile red staining and spinning-disk microscopy imaging

Currently, a variety of environment-sensitive probes are available. We selected Nile Red for several reasons: it offers distinct advantages, including no wash steps, the ability to target a broad range of lipid membranes with environmental sensitivity, and ease of use. Other dyes, such as BODIPY[21], show less environmental sensitivity. Some probes, specifically designed to study the polarity of cell membranes, require more complex labeling procedures, such as Laurdan and Prodan[22]. Nile Red-based membrane probes, such as NR12S[23], have been reported to exhibit higher quantum yields compared to Nile Red; however, these probes are currently only available in the laboratory.

The live cells are stained by Nile Red and imaged by the two sets of spinning-disk microscopes (for the purpose of checking the reproducibility, generalizability, and transfer learning potential), both with doubled spatial resolution of the diffraction limit (evaluated at ~143 nm and ~177 nm, respectively, Figs. S2b and S3d) and at high temporal resolution. The confocal setup can effectively reject the out-of-focus fluorescence signal, which favors the high SNR for quantitative analysis of the improved spectrum ratios. A 473 or 488 nm laser excites the dye, and two emission channels (617/73 nm and 685/40 nm used in this work) are used to collect the fluorescent intensities from the two bands. The resolutions of two microscopes are characterized in Fig. S2b and the detail configurations of microscope 1 (M1) and microscope 2 (M2) are shown in Supplementary Table 3.

Here, we intentionally applied two systems to validate the reproducibility of the DCNN networks, at different excitation, magnification, field-of-view, and detection conditions. We find that the spectral ratio readouts of Nile Red on different imaging modalities, including confocal and LiveSR, remain little changed (Fig. S3a–c). Through our examination of various specimens and imaging conditions—including different dye concentrations, laser intensities, and cell lines (Fig. S4 and Fig. S5)—we demonstrate that our technique exhibits high robustness. We conclude that the emission ratio remains consistent regardless of the change of cell line, illumination intensity, and exposure time employed in our method. The emission ratio may be affected by dye concentrations but only when exceeding the recommended concentration by two-folds, which is unlikely to happen in the experiment. These results conclude that the spectral ratio serves as an intrinsic "optical fingerprint" for reliable and reproducible segmentation of organelle types.

### The inputs of DCNN networks and ground truth

The inputs of the DCNN networks include both the intensity images, by averaging the intensities of the yellow and red channels and being depicted by the brightness, and the red-to-yellow ratiometric images depicted by the pseudo-color, as shown in Fig. 1a. Ratiometric images, in reflection of the lipid polarity of organelles' membranes, have been
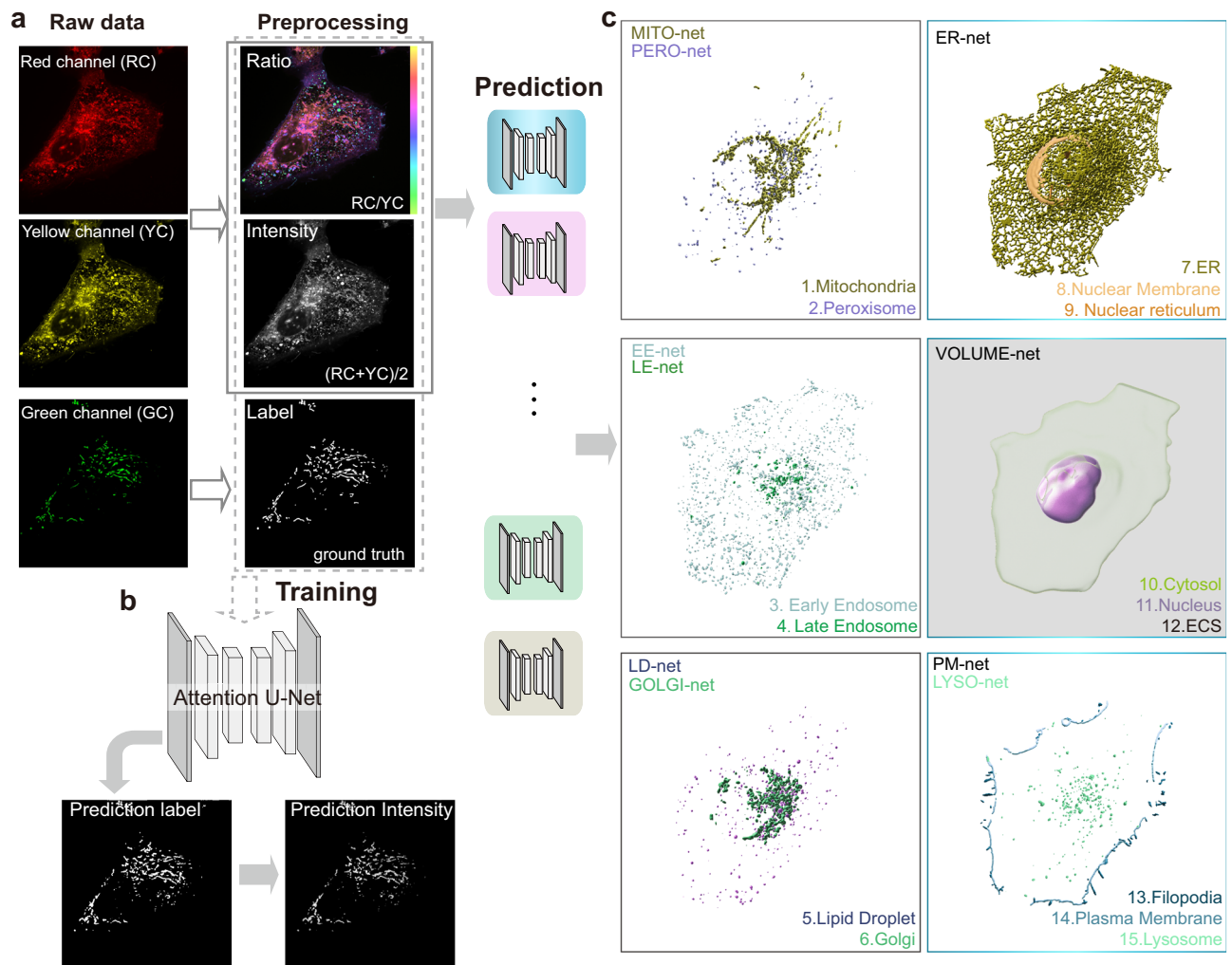
**Fig. 1 | Fast imaging of live cell anatomy of 15 intracellular structures by the universal lipid staining and deep learning segmentation using a spinning disk confocal microscope with an extended resolution. a** Schematic diagram of data acquisition and preprocessing. Extended-resolution image of U2-OS cells labeled by both the organelle-specific GFP and Nile Red are imaged with the same excitation 488 nm and three detection channels (green, yellow, and red). The spectral ratio and intensity are calculated by the yellow channel (Em: 580–653 nm) and the red channel (Em: 665–705 nm). The ground truth label is the binary mask of green channel (Em: 500–550 nm). **b** The untrained Attention U-Net was fed with the ratio, intensity, and label image to train neural networks for distinct organelles. Upon completion of the training process, the model will output the prediction label which will multiply the intensity image to get the prediction intensity. **c** From the super-resolution intensity image and the spectral ratio image, the deep convolutional neural networks (DCNN) predict binary masks of 15 subcellular structures. The masks of cytosol, nucleus, and extracellular space are segmented by VOLUME-net. The masks of ER, nuclear reticulum, and nuclear membrane are segmented by ER-net. The Golgi mask is segmented by GOLGI-net and the lipid droplet mask is segmented by LD-net. Mitochondria mask segmented by MITO-net and lysosome mask segmented by LYSO-net. Early endosome mask segmented by EE-net and late endosome mask segmented by LE-net. Peroxisome mask segmented by PERO-net and the masks of the plasma membrane and filopodia segmented by PM-net (Supplementary Table 1).

found to enhance the segmentation accuracy of the DCNN network, which was particularly effective in discriminating organelles with analogous morphologies. Once the networks are trained, the combined use of both single and multi-segmentation networks can segment different types of organelles by using the ratiometric and intensity images (Fig. 1b–c). To obtain the ground truth masks of different organelles, we first acquire their ground truth images by using GFP plasmids to specifically label a type of organelle each time (online methods). The colocalization channel (green) and the ratiometric intensity channel (yellow and red) are designed with minimal spectrum crosstalk, confirmed in Figs. S6 and S7. The cells labeled by both the organelle-specific GFP and Nile Red are imaged with the same excitation and three detection channels (green, yellow, and red), so that an FG-net segments the foreground area of the ground truth image for the ground truth mask image (Fig. 1a).

With conventional thresholding-based segmentation (Fig. S9d–e), 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 (Fig. S9f–g), which introduces more adjustable parameters. To solve this problem, an FG-net (Fig. S8b) 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. S9c). FG-net provides a more objective and parameter-free approach for precise foreground segmentation. Furthermore, the masks for nuclear membrane (NM), nuclear reticulum (NR), plasma membrane (PM), filopodia (FP), nucleus, cytosol, and extracellular space (ECS), are manually annotated on the high-resolution images (Fig. S10) (online methods).

## Segmentation of intracellular compartments with DCNN networks

Figure 1c shows the segmented structures, including cytosol, nucleus, and ECS, giving the outline of the cells, endoplasmic reticulum (ER), NM, and NR, the reticules membranous structures and the largest membranous system in the cell, Golgi apparatus, lipid droplet (LD), MITO, lysosome (LYSO), early endosome (EE), late endosome (LE), and peroxisome, the major types of organelles, and PM and FP, the outlier membranous structures of the cell. Here, we trained 10 DCNN networks to segment 15 intracellular compartments (Fig. 1, supplementary Movie 1). Namely, LD-net is responsible for recognizing LD, GOLGI-net for Golgi apparatus, MITO-net for MITO, PERO-net for peroxisome, EE-net for EE, LE-net for LE, LYSO-net for LYSO. ER-net is used to recognize the three structures of ER, NM, and NR. PM-net is used to recognize PM and FP, and VOLUME-net is to recognize other three structures of nucleus, cytosol, and ECS.

The ratiometric images and ground truth masks are trained with the networks based on the attention U-Net architecture[24,25] (Fig. S8a). U-Net consists of a U-shaped encoder and decoder, which was designed for biomedical image segmentation[21]. The spatial attention gates are incorporated into the U-Net architecture which is particularly useful on identification and localization of small objects[22]. 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. The selection of activation and loss functions hinges upon whether one is constructing a binary or multi-class classification model. Our setup of decaying learning rate scheduler can effectively prevent overfitting. The learning rate will be reduced to one-third if the validation loss remains no decrease in 5 epochs. The training process will be stopped if the learning rate is less than $1e^{-6}$. The loss, scores, and learning rate curves shown in Fig. S11 and Fig. S12 demonstrate that no overfitting happens during the training. Both the training loss and validation loss keep decreasing during the training process.

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. 2). Compared with the prediction of multiple intensity levels, binary segmentation is simpler and more robust because only true/false decisions are made (Fig. S13). Moreover, the intensities of different colocalization probes vary across cells due to varying expression levels of GFP, while the fluorescent intensities of Nile Red are rather constant. Therefore, 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. S14). These information makes it possible to generate the digital anatomy and dynamics of live cells.

During training, we 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 0.9 to 1.1×. The intensity images are randomly multiplied by 0.75 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 empirically assign different class weights for different DCNN networks. The class weights are hyperparameter of the Softmax loss, which is used to address class imbalance in classification problems. 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 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. 2 for the 3D datasets, the merged images and their quantitative analysis show that the overall pixel accuracies of over 91.7% have been achieved for all the intracellular structures (Fig. 3a). We further evaluate the performance of the DCNN segmentation in Fig. 3b using F1scores, recall, and precision to reduce the bias of pixel accuracies. Again, the binary segmentation networks resulted in the high F1scores for Golgi apparatus (0.66), LD (0.72), MITO (0.63), peroxisomes (0.68), EE (0.59), LE (0.51), and LYSO (0.67) (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 and produce the multiplexing images of subcellular compartments from the ratiometric images, we further evaluate the performance of the "in-silico" multi-(pseudo) color images, referring specifically to the DCNN-predicted multi-(pseudo) color outputs. The PCC and the Manders' overlapping coefficient (MOC)[26] 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, LD, 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). To provide a comprehensive evaluation of our training accuracy, we additionally assessed the PCC and Dice score for five organelles in Fig. 3h–i. 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 the organelle segmentation results using other imaging modalities, our results show superior or comparable performance. For example, the predicted fluorescence images of Golgi apparatus on transmitted light images[12] only reaches a low-PCC value of less than 0.2. The Golgi segmented by our method reaches the PCC value of 0.77. Although predicted fluorescent images from the bright field images can achieve PCC values of -0.7 for MITO, the ultrafine structures such as the cristae cannot be resolved[12]. Compared with organelle segmentation from electron microscopy dataset[3] collected using focused ion beam scanning electron microscopy, our results achieve a slightly worse performance on ER and MITO, but an evident superior performance on vesicle organelles, such as LYSO (improved from 0.53 to 0.67), and endosomes (improved from 0.48 to 0.55).

We further compared the segmentation accuracy of ratio net and intensity net to assess the impact of incorporating spectral ratios. The results indicate that the spectral ratio images play a critical role in enhancing the segmentation accuracy for vesicle organelles with similar morphology (Fig. S15). For example, the Ratio-net improves the F1 score than intensity-net in the dataset of EE (0.59 vs 0.43, Fig. S15e–h) and Golgi apparatus (0.66 vs 0.54, Fig. S15i–l). For structures with distinct morphologies, such as MITO, the ratio image seems to add little value, as the F1scores are similar for ratio-net and intensity-net (0.63 vs 0.62, Fig. S15a–d). Interestingly, the ratio-net and intensity-net also achieve similar F1 scores for LD (0.72 vs 0.69). It is possibly due to the fact that LD are the brightest organelles after Nile Red staining.

## The organelles' segmentation results across different instruments and cell types

We further evaluate the generalizability power of the DCNN networks. For the generalization test, we firstly train the DCNN networks with the images of five different organelles acquired on the microscope M1, and test the performance of DCNN networks on the images acquired from a different microscope M2, including different sets of excitation wavelengths, objectives, camera sensors, spinning disk systems, and system resolutions (Supplementary Table 3). The MOCs of ER, MITO,
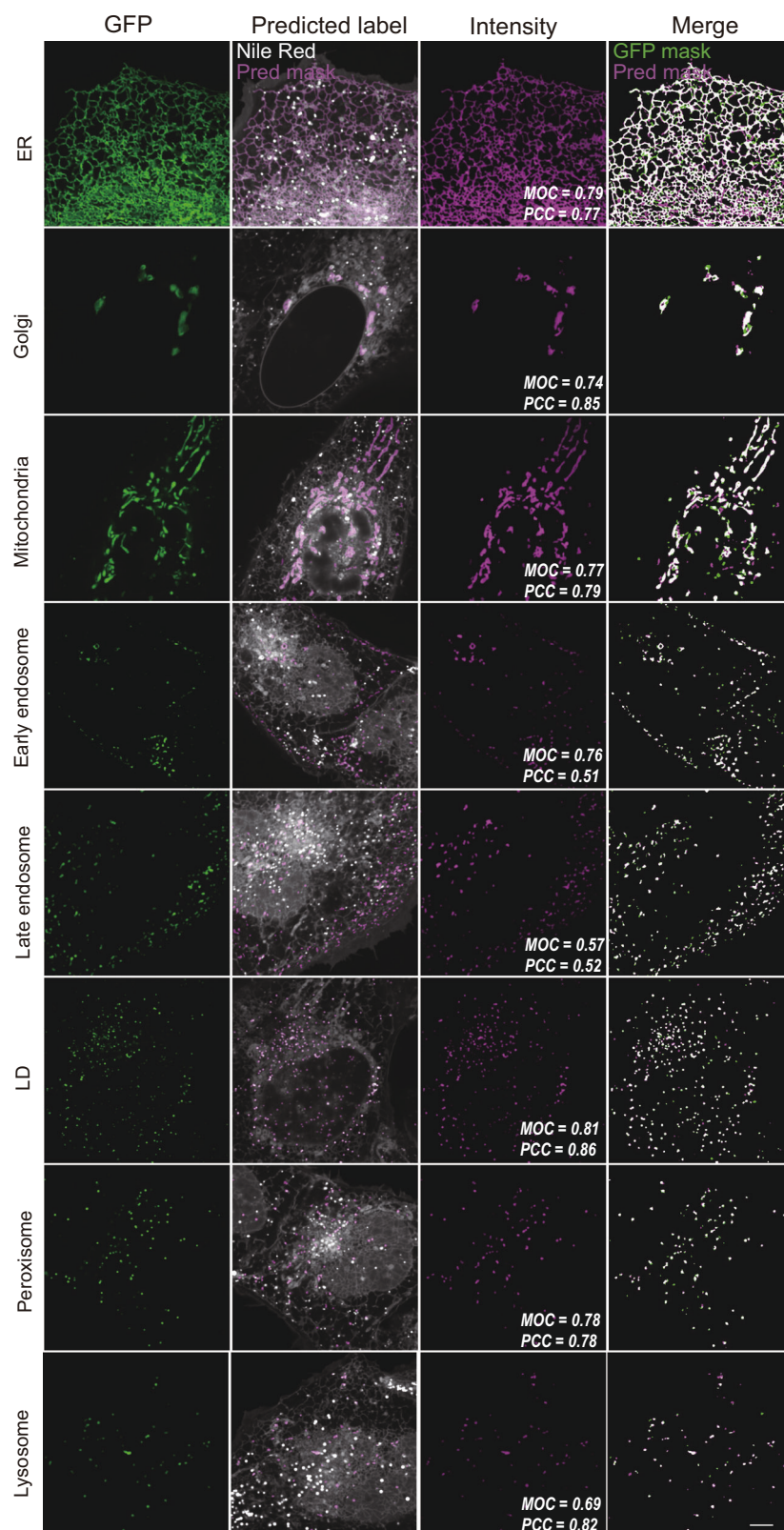
**Fig. 2 | Accurate in-silico prediction of the eight intracellular structures based on the GFP ground truth images.** (the first column). The second column shows the Nile Red stained fluorescence images with magenta masks predicted by the networks, the third column shows the intensity images by multiplying the fluorescence image and the predicted masks, and the fourth column shows the merged images of predicted masks (in magenta) and ground truth masks (in green, FG-net prediction of ground truth intensity images). In the merged images, white or black pixels indicate true prediction, magenta pixels indicate false positive, and green pixels indicate false negative. For the other seven manually annotated structures see Supplementary Fig. S10. All independent experiments were replicated at least three times ($n \geq 3$). MOC Manders' overlapping coefficient. PCC Pearson's correlation coefficient. Scale bar: 5 μm.
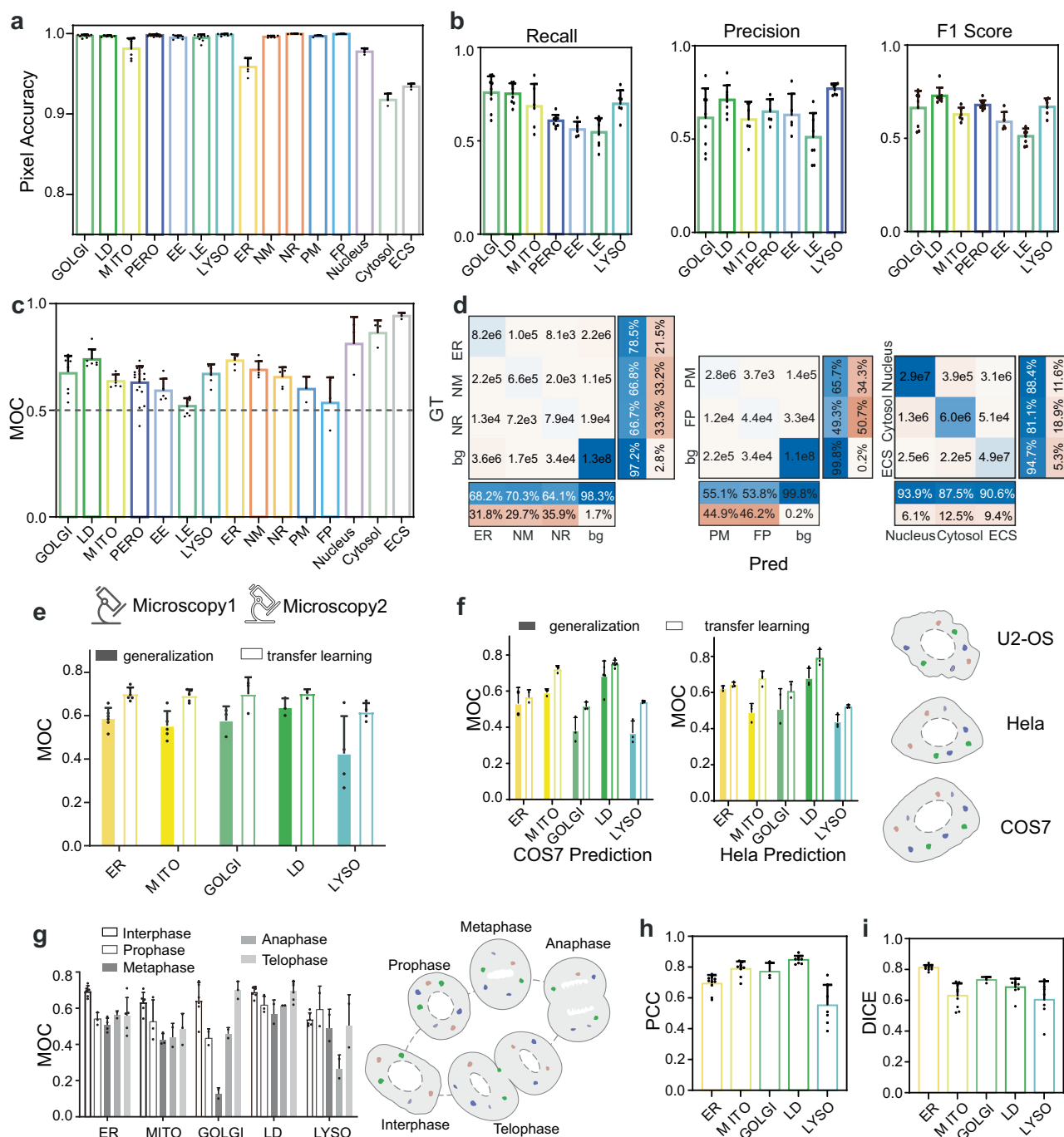
**Fig. 3 | Comprehensive evaluation of the segmentation accuracies for the 15 cellular structures. a** Pixel accuracies of the 15 structures. **b** Recall, precision, and f1 score of binary segmentation networks. **c** The Manders' overlapping coefficient (MOC) of 15 structures between predicted masks and ground-truth masks. **d** The confusion matrixes of multi-classification networks including ER-net, PM-net, and VOLUME-net. The DCNN networks in (**a**–**d**) are trained and evaluated from the dataset acquired from microscope M1. **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 micro-scope 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 transfer 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. **h** The Pearson Correlation Coefficient (PCC) of different organelles between predicted intensity and GFP intensity. **i** The Dice score of different organelles' prediction accuracy. Error bars represent the mean values ± standard deviation (S.D.) of indi-cated independent experiments.

and LD still show good colocalization results, while the MOCs of Golgi and LYSO show weak colocalization (Fig. 3e, generalization). More-over, we train the DCNN networks with U2-OS datasets and use them to predict the COS7 (Fig. 4d–e, Fig. S16) and Hela datasets. The MOCs of ER, MITO, Golgi, and LD show good colocalization results, while the

MOC of LYSO shows weak colocalization (Fig. 3f, generalization). In short, most of DCNN networks show satisfactory generalizability across different microscopes and cell lines. Only some of them require transfer learning with additional data, as after transfer learning with additional training dataset acquired on M2, the MOCs of all structures
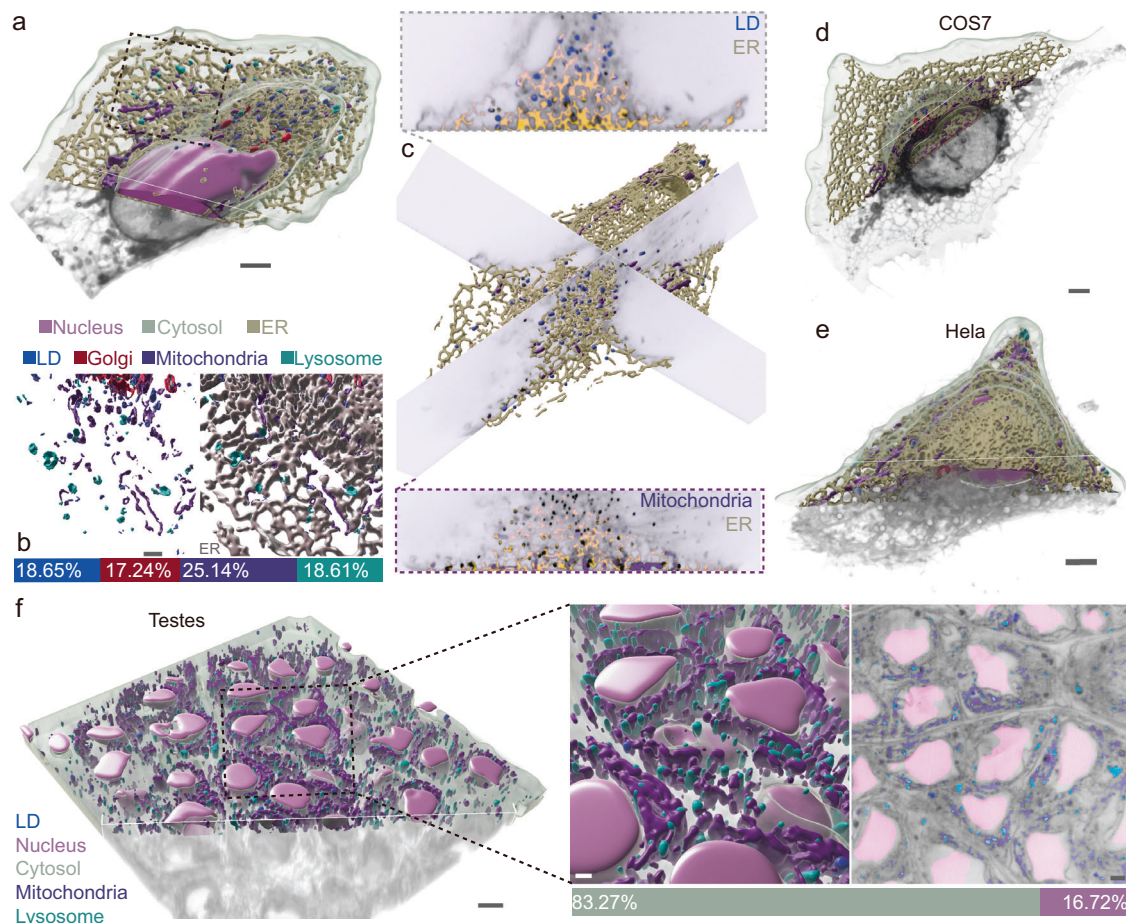
**Fig. 4 | Organelle segmentation on different biology samples by the universal lipid staining and deep learning. a** Segmentation results of U2-OS cells. **b** Zoomed-in view in the black box in (**a**), which shows interactions of lipid droplet, Golgi, Mitochondria, and lysosome with ER. The bar shows 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 3$^{rd}$ instar larvae testes in *Drosophila* by the universal lipid staining and deep learning. Scale bar: 5 μm.

show good colocalization (MOC > 0.5) with the ground truth masks (Fig. 3e, transfer learning). With only 5–6 additional images of COS7 and Hela images, the prediction accuracies became significantly improved (Fig. 3f, transfer learning). These results underscore the transfer learning capabilities of our method. For biologists aiming to investigate organelle interactions under specific cellular states or experimental conditions, the addition of a minimal amount of ground truth data is recommended to enable the network to produce robust predictive results, thereby facilitating further biological discoveries.

Moreover, we demonstrate that our approach can be applied to image the anatomic structures of living tissues. As cells in living tissues show distinct morphologies with the cultured cells, the DCNN networks need to be re-trained and evaluated independently. We show that Nile Red can homogeneously stain the membranous structures through the 3$^{rd}$ instar larvae testes (-100 μm thick), and in contrast, organic dyes, such as Mitotracker, can hardly penetrate cyst cells to diffuse homogeneously into spermatogonia (Fig. S1h). By conducting the colocalization experiments to recognize LD, LYSO (LysoView), and MITO (Mito-moxMaple3), and manually labeling the nucleus, cytosol, and ECS, as ground truth, we have successfully realized the six-color imaging of the intracellular organelles of complex multi-cellular sample of the entire testes system (Fig. 4f). The nucleus, cytosol, and ECS are relatively simple structures but challenging to colocalize using specific labeling methods in larvae. Therefore, we manually label these structures. This demonstration opens the door to study the intracellular organelles' interactions in living animal and across intracellular networks in tissues.

## Cell anatomy imaging across mitotic stages

According to the organization of intracellular structures, such as chromosomes, cytoskeleton, etc., the mitotic cell division experiences four discrete phases: prophase, metaphase, anaphase, and telophase[27,28]. As the spatiotemporal rearrangements of membranous compartments are crucial to the propagation of cells, based on the spinning disk microscopy results with an extended spatial resolution, here we use the in-silico multi-color imaging to study the 3D anatomy of live U2-OS cells across mitotic phases. We demonstrate that the DCNN networks can perform accurate segmentations for most organelles during mitosis (Fig. 3g). As shown in Fig. 5a, b and Fig. S17, the nuclei are clear in interphase, prophase, and telophase, become less obvious in anaphase, and almost disappear in metaphase. The NR can be resolved in interphase and prophase, as it promotes the disassembly of NM[28]. In contrast, the PM becomes thicker and brighter since metaphase, which reflects their mechanical property changes (Fig. 5g). Interestingly, our results can further quantitatively reveal the relatively small changes in the volume of the cells across different phases, ranging from 4000 to 6000 μm$^3$ (Fig. 5h, Fig. S14a).

ER is the largest membranous organelle within the cell. The results are presented in Fig. 5c and Figs. S17b and S18 show a denser ER structure around the nucleus with a mixture of tubule and sheet

|  | Interphase | Prophase | Metaphase | Anaphase | Telophase |

a, b, c, d, e, f, g, h, i

Cytosol / Nucleus / ECS

ER / Nuclear Membrane / Nuclear reticulum

Golgi / LD

Mito / Lyso

Early Endosome / Late Endosome

Filopodia / Peroxisome / Plasma Membrane

cytosol  nucleus   ER  Golgi  LD   Mito  Lyso  EE  LE  Pero  NM  NR  PM  FP

structures in interphase, while it has more tubule structures uniformly distributed in metaphase and anaphase. Figure 5d, Figs. S14c and S18 accordingly show a 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 extended-resolution in-silico images in Fig. 5e and Fig. S18 further

reveal the elongated and stretched MITO in interphase and prophase compared with the fragmented structure in metaphase ready for equal inheritance[29]. The total volume analysis of MITO shows the decrease from interphase to mitosis.

The 3D digital cell anatomy, as shown in Fig. 5 and supplementary Movies 2–6, further allows a statistical analysis of all the intracellular structures. For example, Fig. S17b plotted the spatial distributions of

**Fig. 5 | Dynamic imaging of cell anatomy across all the mitosis stages.**
**a** Extended-resolution images 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. S14 for detailed volume statistics of different intracellular structures. Each independent experiment representing distinct mitotic stages was replicated at least three times ($n \geq 3$). Scale bar: 5 μm.

different compartmental pixels to the coverslip or nucleus, and Figs. S14 and S19 showed the particle volume, number, sphericity, and emission ratio of vesicular organelles, including LD, LYSO, EE, LE, and LYSO in different mitotic stages. Notably, among these results LD data analytics is also consistent to the early observation using the conventional sophisticated fluorescence labeling that the nucleoplasmic LDs disappear during mitosis and form anew in daughter cells in the MTP-dependent manner[30].

### Fast imaging of the organelles' interactome
Next, we apply the deep learning enhanced extended-resolution segmentation of membranous structures to study the intracellular organelles' interactome, as the organelles coordinate with each other and undergo spatiotemporal interactions to perform cellular functions. By precisely segmenting organelles, we can observe the interactions between different organelles. Figure 4a–c depicts the interactions between LD, Golgi, MITO, and LYSO with ER in U2-OS cells. Figure 4b extracts the contact sites between the ER and these different organelles and provides a statistical analysis of the proportion of interaction sites. From the cross-sectional view in Fig. 4c, we can clearly observe the interactions between ER and LD as well as ER and MITO.

As the speed of 3D volumetric imaging is insufficient in capturing the fast dynamic interactions, we turn to the 2D image acquisition by focusing on an arbitrary optical section of up to five sections within a region 1.6 μm away from the coverslip, which provides consistent and reliable dataset to train the DCNN networks (Fig. S20). Here, we can typically achieve a high throughput of nine structures per exposure, including the eight types of membrane organelles stained with Nile Red and microtubule structures labeled by tubulin GFP, where the three detection channels of two Nile Red ratio channels and one GFP channel are projected onto the separate areas of a sCMOS camera via an image splitter (Fig. S2a). Notably, as the networks' performance based the 2D image data is less satisfactory for peroxisome and LE (medium colocalization of MOC > 0.3) and for the EE (weak colocalization of MOC = 0.259), we only count here the five sets of predictions for LD, ER, Golgi apparatus, MITO, and LYSO (strong or good colocalization of MOC > 0.5), therefore the reliable high throughput of six structures per exposure, including the five types of membrane organelles and the microtubule structure, are presented. In comparison, the linear unmixing method based on spectral detection requires 3 lasers and 26 images to solve the 6 structures[1], that is, a throughput of 0.23 color/structure per exposure, and the linear unmixing method based on the six channels of time-modulated laser excitations requires 6 or 8 image exposures to resolve the six structures[1,4], resulting in a throughput calculated as 1.0/0.75 color per exposure. Collectively, in-silico multiplexing imaging, presented in our work, has increased the imaging throughput by an order of 6–26 folds.

With both the original fluorescent images of microtubule and Nile Red stained images of sub-cellular membrane compartments (Fig. 6a), the unknown "big picture" of multiple intracellular organelles' interactions can be revealed, using Nile Red staining and DCNN networks. In the enlarged four-color image series of microtubule, LD, MITO, and LYSO (Fig. 6c), the fission and fusion process of MITO, the interactions between LD and microtubule, and the interactions among MITO, LYSO, and microtubule are presented. Figure 6d clearly shows the moment when the Golgi vesicles fuse and interact with microtubule. Figure 6e records the process of ER interactions with LD and LYSO. Supplementary Movie 7 shows the bird view of the sophisticated multiple interactions among the six subcellular structures, which series of data images were simultaneously acquired in high speed and high throughput.

## Discussion
The conventional fluorescence imaging techniques suffer from the issues of low throughput, insufficient speed, or low resolution for live cell organelle interactome. This has been limited by the rooted fact associated with "one-to-one" specific labeling strategy, including the availability of fluorescent colors, the spectrum cross-talks, labeling efficiency and throughput issues, etc. The signal level of each color channel must be sacrificed to achieve multispectral unmixing, which requires a longer image acquisition time and efforts in switching between excitation and detection channels, which often causes motion blur. Although SR techniques have been well developed for decades towards high spatial resolutions, it is often at the price of temporal resolution. Here, we presented a different strategy, that is, avoid using specific labeling but use a generic lipid dye to stain all the up to 15 membrane organelle structures. As the dye presents excellent chromatic polarity sensitivity to the membrane environment, ratiometric imaging at the two spectral bands (617 nm/685 nm) will reveal the different situations (e.g., polar or nonpolar) of the dye inside the different organelles. Furthermore, the use of a commercial spinning-disk confocal microscopy system with extended resolutions in both lateral and axial axes can effectively reject the out-of-focus fluorescence with a high SNR for ratiometric analysis, which set of data, as comprehensive optical fingerprints of organelles, enable DCNN to recognize and segment organelles. Moreover, we have adapted the method to super resolution structured illumination microscopy (SR-SIM) successfully to demonstrate its compatibility (Fig. S24). This work has provided an approach for fast and robust segmentation and multiplexing imaging of multiple organelles in live cells.

### Limitations of our technique
The use of DCNN networks always causes concerns over their accuracies. This limitation partly arises from imperfect ground truth data. For example, in our current work, to obtain the training data, MITO was labeled using GFP, as MitoTracker Green introduces spectral interference with NileRed, but GFP labels the entire mitochondrial matrix, beyond the mitochondrial membrane, which resulted in the look of the whole mitochondrial structures. Moreover, it is crucial to recognize that the labeling positions for GFP and NileRed exhibit some degree of non-congruence. As a result of this disparity, the linearity of the signal has been notably compromised. Therefore, the development of a more precise ground truth labeling approach to precisely mark membrane structures of various organelles will significantly enhance the precision of deep learning segmentation.

Furthermore, except the use of binary masks instead of intensity images to avoid the generation of fake signals, the strategies of using a variety of specimens and imaging data from different microscopes, from different cell lines, and on different phases of cell cycles can improve the robustness of the DCNN networks. In experiments where
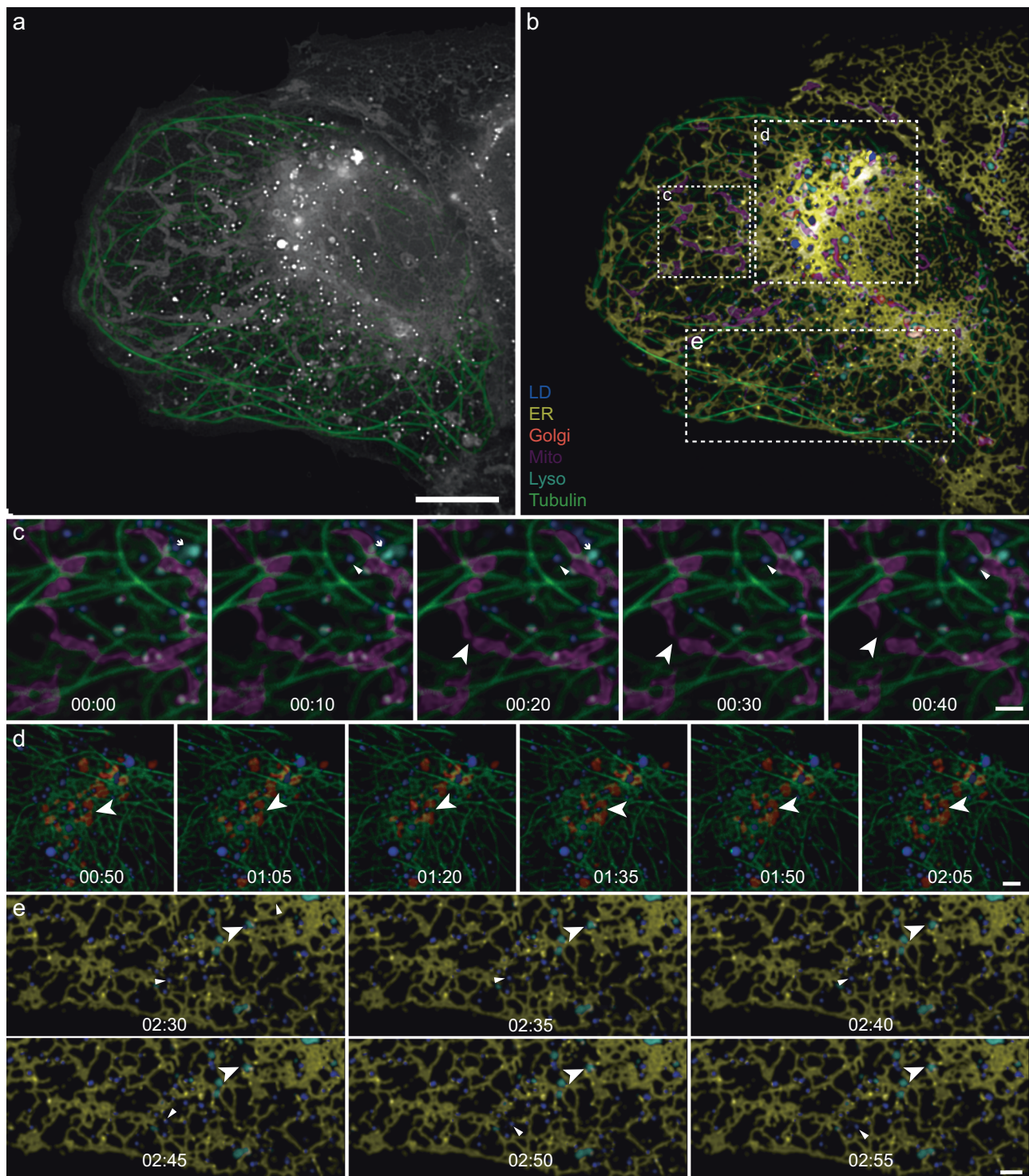
**Fig. 6 | Fast and multiplexing imaging of the dynamic interactions among multiple organelles in a live cell. a** The extended-resolution images of microtubule (green) and membranous structures (gray). **b** Different types of organelles are segmented and the masks are applied to the intensity image, resulting in the multi-color image of lipid droplet, ER, Golgi apparatus, mitochondria (Mito), lysosome (Lyso), and microtubule (Tubulin). (Fig. S21) **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 lipid droplet (Fig. S22). **e** Zoom-in time-lapse images of the interactions among ER, lipid droplet, and lysosome. Three independent replicates were conducted. Scale bar: (**a**, **b**) 10 µm, (**c**–**e**) 2 µm.

organelle morphology and lipid membrane composition undergo minimal changes, the existing network can be directly applied for predictions. If the accuracy prove insufficient, additional training data can be incorporated to enhance performance. We have demonstrated that, under specific cellular conditions or experimental interventions, such as drug treatment or cellular stress, the accuracy of direct predictions is comparable to that achieved with fintuned models (Fig. S29 and Fig. S30). Careful evaluations of the error rates of the predicted results remain essential before meaningful biological conclusions can be drawn. We show that the MOC/F1scores, typically in the ranges of

0.5 to 0.7 for 3D datasets and 0.5 to 0.6 for the five organelles on 2D datasets, varying across different cellular compartments, are within the confidence. In densely intracellular regions, especially near the nucleus with high organelle concentration, segmentation accuracy may decrease. High-resolution imaging with enhanced ability for background removal is essential to obtain high-quality raw images. For the low-fidelity results, cross-validation experiments are still needed. Besides, the networks should be fine-tuned with additional training datasets under similar conditions in the experiments when using the pre-trained DCNN networks.

## Outlook of our technique

With all the data acquired from the commercial microscopes, only one fluorescent stain, one excitation laser, and two color detections have been found sufficient in resolving the digital anatomy of the 15 intracellular compartments and their dynamic interactions. Our method fundamentally overcomes the limitations of color crosstalk and the low joint labeling efficiency of multiple probes in the spectral domain. It significantly increases both the imaging speed and throughput without photobleaching concerns in live-cell imaging. Our approach and demonstrations suggest an opportunity in using the denoising power of deep learning[31,32] for long-term observation of cells, and the multi-color imaging of complex live tissues over a large field of view (Fig. S23).

While lipid dyes have gained considerable interest in subcellular imaging[20,33–36], the specificity of lipid dyes becomes a major challenge, due to the large variety of lipid molecules and membrane structures that exist in a cell. Our method of deep learning segmentation gears the disadvantage of unspecific staining into an advantage to enable multiplexing 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 imaging of the physicochemical properties of subcellular membranes. Our work further suggests that a variety of recently developed lipid dyes[23,34,36,37] should be compatible with our method. Taking these together, the features, provided by lipid dyes with high brightness and biocompatibility to emit responsive emissions across the heterogeneous organelle structures, are highly favorable to improve the segmentation accuracies in resolving the organelles with similar morphology.

Here, we have extended the use of lipid dyes in SR-SIM, and the results presented in supplementary Fig. S24 indicate that the higher resolution of ~110 nm offered by a super resolution SIM can improve the accuracy to resolve the cristae of MITO stained by NileRed. Deep learning segmentation of organelles can also be achieved in a broad range of SR microscopes, such as stimulated emission depletion[33] and single molecule localization microscopy[23,35,38], though the suitability in large field of view live cell imaging yet to be evaluated. Moreover, to reduce phototoxicity, adoption of our method to long-term volumetric imaging techniques, such as lattice light sheet[39], may lead to the tracking of organelle dynamics across mitotic stages of cells in their native environment.

Although the current technique focuses on imaging membranous structures, our approach can be supplemented by using strategies to target the membrane-less organelles, such as ribosomes, centrosomes, and cytoskeleton, for simultaneous imaging of more than 15 organelles. As our method is orthogonal to other existing approaches, such as multi-color labeling and spectrum unmixing, so that the level of complexity can be further increased by combing other fluorescent probes to study the subcellular interactions among organelles, cytoskeleton, and single molecular proteins. Besides the fluorescence intensity and wavelength, the fluorescence polarization[20] and lifetime[40,41] 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% $CO_2$. 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. The 1 μg/ml Nile Red concentration was tested through pre-experiment and chosen based on considering both the signal-to-noise ratio of fluorescent intensity (if too low) and the phototoxicity for long-term imaging (if too high).

For colocalization experiments, the cells were transfected 16 h before imaging with the plasmids of EE-GFP (Rab5a, C10586, BacMam 2.0, CellLight, USA), LE-GFP (Rab7a, C10588), ER-GFP(ER signal sequence of calreticulin and KDEL, C10590), Golgi-GFP (Golgi-resident enzyme N-acetylgalactosaminyltransferase 2, C10592), MITO-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 LYSO is obtained by Lysoview™ 488 (70067-T, Biotium) 30 min before imaging and without washing during imaging. For mitosis experiments, the nuclei ware 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 min, the testes stained with Nile Red can be imaged under fluorescent microscopy. For imaging of LYSO and LD, we used wild-type *Drosophila* (w1118). For colocalization imaging of MITO, we used *bam*-Gal4 to drive UASp-*mitochondria*-Moxmaple3 expression in male germ cells.

The comparison of traditional multicolor imaging using different cells under the same experiment protocol (Figs. S25–S28). We performed the comparative experiment using different wells in an μ-Slide 8 Well (80827, ibidi). The cells were from the same batch. For the control group, cells were stained with NileRed 30 min before imaging. For the experimental group, cells were transfected with ER (GFP-KDEL) and Golgi (mCherry-Golgi) plasmids using Lipo3000 24 h prior to imaging, and stained with commercial HBmito Crimson (https://www.medchemexpress.cn/hb-imm-d-red.html) for MITO 30 min before imaging.

### Spinning-disk microscopy

Two microscopes which are slightly different are used in our experiments. M1: an inverted fluorescence microscope (TI-E, Nikon) is equipped with a TIRF objective (CFI Apochromat TIRF ×100 oil, NA 1.49, Nikon), and a spinning disk confocal system (CSU-W1, Yokogawa). The SR imaging module (Live SR, Gataca) can 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: S525/50 m; yellow: S617/73 m; red: S685/40 m; Chroma) in the spinning disk system. M2: an inverted fluorescence microscope (TI-2, Nikon) equipped with a TIRF objective (CFI Apochromat TIRF ×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 SR imaging module (Live SR, Gataca) could double the imaging resolution. Emission

fluorescence passes through a multi-splitter module (CAIRN, Multi-Split V2) with three dichroic mirrors (ZT488rdc/ZT561rdc/ZT640rdc; Chroma) and different emission filter (blue: ET480/40 m; green: S525/50 m; yellow: S617/73 m; red: S685/40 m; Chroma). Then, illuminated on the different parts of the camera (Prime BSI, Photometrics). Most of the datasets were acquired from M2. Half the 3D datasets of Golgi apparatus, LD, MITO, and ER are acquired from M1. 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). The detailed configuration of M1 and M2 is shown in Supplementary Table 3.

### Structure illumination microscopy

The training, validation, and test data in Fig. S24 are collected through the Polar SIM system (Airy Technology Co., Ltd., China). A 488 nm laser is used to excite Mito Tracker Green, and a 561 nm laser is used to excite NileRed. Three emission channels (green: ET525/50 m; yellow: ET595/44 m; red: ET706/95 m; Chroma) were detected for ground truth and ratiometric (red/yellow) imaging. The SIM reconstruction process was conducted using the Airy-SIM software and Open-3DSIM platform[42].

### Training data preparation

Multi-channel images are registered with chromatic aberration correction algorithm[43]. 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 the yellow channel. For the 3D dataset, the whole imaging FoV is 1200*1200*Nz px$^3$ (Nz ranges from 12 to 30), which is resized into 802*802*24 px$^3$ and is further cropped into random 256*256*24 px$^3$ patches. For the 2D dataset, the images are resized and cropped into random 256*256 px$^2$ patches.

For the structures of the NM, NR, PM, FP, nucleus, cytosol, and ECS, the ground truth masks are manually annotated. The annotation is performed with Image Labeler app from MATLAB. The NM and NR are labeled on the ER-GFP image since ER-localized proteins simultaneously localize to ER, NM, and NR. The PM, FP, nucleus, cytosol, and ECS are labeled on the Nile Red fluorescent images through their morphology recognized by biologists. For other structures, the green-channel ground truth 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*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 ground truth image by a pre-trained FG-net (Fig. S8b), which is trained with 40 manually annotated ground truth 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. S8b), and the subcellular segmentation networks take 3D intensity and ratio images as input (2@256*256*24) (Fig. S8a). 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 MITO than as peroxisome, so the weight of MITO 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 are slightly different (Supplementary Table 2).

### Image analysis and display

The evaluation of predicted masks is performed by custom-written MATLAB. See Supplementary Note for specific assessment and calculation methods. 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 (Figs. S14, S17, S19) 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. S19). The spot tool in Imaris was used to find the center of the object to calculate the distance between the nucleus and the coverslip (Fig. S17).

### Reporting summary

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

## Data availability

All the data are available in the main text or the supplementary materials. We open-sourced the SIM imaging dataset on Figshare: https://figshare.com/s/796baf8a3f4776054940. The datasets contain 24 high-quality SIM images and corresponding 3D wide-field imaging for training, with paired GFP-labeled organelles (mitochondria) as ground truth.

## Code availability

Our PyTorch implementation of the DCNN networks is publicly available at https://github.com/KarlZhanghao/LiCAI, which contains the pretrained models and the prediction of 15 structures of cells in Fig. 1 and Supplementary Movie 1. The model training and transfer learning examples of representative organelles along with the training dataset are within the same repository. A demonstration of the model in the BioImage format can also be found at https://doi.org/10.5281/zenodo.7653695.

## References

1.   Valm, A. M. et al. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017).
2.   Cusick, M. E., Klitgord, N., Vidal, M. & Hill, D. E. Interactome: gateway into systems biology. *Hum. Mol. Genet.* **14**, R171–R181 (2005).
3.   Heinrich, L. et al. Whole-cell organelle segmentation in volume electron microscopy. *Nature* **599**, 141–146 (2021).
4.   Chen, K., Yan, R., Xiang, L. M. & Xu, K. Excitation spectral microscopy for highly multiplexed fluorescence imaging and quantitative biosensing. *Light Sci. Appl.* **10**, 97 (2021).
5.   Nixon-Abell, J. et al. Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. *Science* **354**, aaf3928 (2016).
6.   Shroff, H., Galbraith, C. G., Galbraith, J. A. & Betzig, E. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nat. Methods* **5**, 417–423 (2008).
7.   Huang, X. et al. Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy. *Nat. Biotechnol.* **36**, 451–459 (2018).
8.   Yang, X. S. et al. Mitochondrial dynamics quantitatively revealed by STED nanoscopy with an enhanced squaraine variant probe. *Nat. Commun.* **11**, 3699 (2020).

9. Liu, T. et al. Multi-color live-cell STED nanoscopy of mitochondria with a gentle inner membrane stain. *Proc. Natl. Acad. Sci.* **119**, e2215799119 (2022).

10. Tyson, J. et al. Extremely bright, Near-IR emitting spontaneously blinking fluorophores enable ratiometric multicolor nanoscopy in live cells. *ACS Cent. Sci.* **7**, 1419–1426 (2021).

11. Huang, B., Bates, M. & Zhuang, X. Super-resolution fluorescence microscopy. *Ann. Rev. Biochem.* **78**, 993–1016 (2009).

12. Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F. & Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat. Methods* **15**, 917–920 (2018).

13. Christiansen, E. M. et al. In silico labeling: predicting fluorescent labels in unlabeled images. *Cell* **173**, 792–803 (2018).

14. Manifold, B., Men, S., Hu, R. & Fu, D. A versatile deep learning architecture for classification and label-free prediction of hyperspectral images. *Nat. Mach. Intell.* **3**, 306–315 (2021).

15. Cheng, S. Y. et al. Single-cell cytometry via multiplexed fluorescence prediction by label-free reflectance microscopy. *Sci. Adv.* **7**, eabe0431 (2021).

16. Guo, Y. et al. Visualizing intracellular organelle and cytoskeletal interactions at nanoscale resolution on millisecond timescales. *Cell* **175**, 1430–1442 (2018).

17. Azuma, T. & Kei, T. Super-resolution spinning-disk confocal microscopy using optical photon reassignment. *Opt. Exp.* **23**, 15003–15011 (2015).

18. Greenspan, P., Mayer, E. P. & Fowler, S. D. Nile red: a selective fluorescent stain for intracellular lipid droplets. *J. Cell Biol.* **100**, 965–973 (1985).

19. Sackett, D. L. & Wolff, J. Nile red as a polarity-sensitive fluorescent probe of hydrophobic protein surfaces. *Anal. Biochem.* **167**, 228–234 (1987).

20. Zhanghao, K. et al. High-dimensional super-resolution imaging reveals heterogeneity and dynamics of subcellular lipid membranes. *Nat. Commun.* **11**, 5890 (2020).

21. Loudet, A. & Burgess, K. BODIPY dyes and their derivatives: syntheses and spectroscopic properties. *Chem. Rev.* **107**, 4891–4932 (2007).

22. Owen, D. M., Rentero, C., Magenau, A., Abu-Siniyeh, A. & Gaus, K. Quantitative imaging of membrane lipid order in cells and organisms. *Nat. Protoc.* **7**, 24–35 (2012).

23. Kucherak, O. A. et al. Switchable nile red-based probe for cholesterol and lipid order at the outer leaflet of biomembranes. *J. Am. Chem. Soc.* **132**, 4907–4916 (2010).

24. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. *Med. Image Comp. Comput.-Assist. Interv.* - MICCAI 2015, **9351**, 234–241 (2015).

25. Oktay, O. et al. Attention U-Net: learning where to look for the pancreas. Preprint at *arXiv* https://arxiv.org/abs/1804.03999 (2018).

26. Manders, E. M. M., Verbeek, F. J. & Aten, J. A. Measurement of co-localization of objects in dual-colour confocal images. *J. Microsc.* **169**, 375–382 (1993).

27. Jongsma, M. L., Berlin, I. & Neefjes, J. On the move: organelle dynamics during mitosis. *Trends Cell Biol.* **25**, 112–124 (2015).

28. Carlton, J. G., Jones, H. & Eggert, U. S. Membrane and organelle dynamics during cell division. *Nat. Rev. Mol. Cell Biol.* **21**, 151–166 (2020).

29. Moore, A. S. et al. Actin cables and comet tails organize mitochondrial networks in mitosis. *Nature* **591**, 659–664 (2021).

30. Sołtysik, K., Ohsaki, Y., Tatematsu, T., Cheng, J. & Fujimoto, T. Nuclear lipid droplets derive from a lipoprotein precursor and regulate phosphatidylcholine synthesis. *Nat. Commun.* **10**, 473 (2019).

31. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).

32. Li, X. Y. et al. Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising. *Nat. Methods* **18**, 1395–1400 (2021).

33. Spahn, C., Grimm, J. B., Lavis, L. D., Lampe, M. & Heilemann, M. Whole-Cell, 3D, and multicolor STED imaging with exchangeable fluorophores. *Nano Lett.* **19**, 500–505 (2019).

34. Collot, M. et al. Probing polarity and heterogeneity of lipid droplets in live cells using a push-pull fluorophore. *Anal. Chem.* **91**, 1928–1935 (2019).

35. Moon, S. et al. Spectrally resolved, functional super-resolution microscopy reveals nanoscale compositional heterogeneity in live-cell membranes. *J. Amer. Chem. Soc.* **139**, 10944–10947 (2017).

36. 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 Adv.* **4**, 8481–8488 (2014).

37. Collot, M. et al. MemBright: a family of fluorescent membrane probes for advanced cellular imaging and neuroscience. *Cell Chem. Biol.* **26**, 600–614.e607 (2019).

38. Sharonov, A. & Hochstrasser, R. M. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proc. Nat. Acad. Sci.* **103**, 18911–18916 (2006).

39. Chen, B. C. et al. Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014).

40. Lu, Y. et al. Tunable lifetime multiplexing using luminescent nanocrystals. *Nat. Photon.* **8**, 32–36 (2013).

41. Castello, M. et al. A robust and versatile platform for image scanning microscopy enabling super-resolution FLIM. *Nat. Methods* **16**, 175–178 (2019).

42. Cao, R. et al. Open–3DSIM: an open-source three-dimensional structured illumination microscopy reconstruction platform. *Nat. Methods* **20**, 1183–1186 (2023).

43. Matsuda, A., Koujin, T., Schermelleh, L., Haraguchi, T. & Hiraoka, Y. High-accuracy correction of 3D chromatic shifts in the age of super-resolution biological imaging using chromagnon. *J. Vis. Exp.* **160**, e60800 (2020).

## Author contributions

D.J., P.X., M.L., and K.Z. conceived and designed the project. D.J., P.X., and M.L. supervised the research. M.L. performed experiments and acquired all experimental data. K.Z. developed and trained the deep learning network. X.C., W.L., and J.W. assisted in optimizing the deep learning network architecture. T.L. prepared *Drosophila* testis samples. Y.W. and Z.W. performed manual data annotation. F.S. conducted a ratio statistical analysis. C.S., Y.Z., and J.F. provided a biological interpretation of the results. M.L. and K. Z. conducted data analysis and visualization. The manuscript was written by M.L., K.Z., P.X., and D.J., with critical feedback from all authors. All authors reviewed and approved the final manuscript.

## Competing interests

P.X. is a consultant of Airy Technology Co., Ltd. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-57877-5.

**Correspondence** and requests for materials should be addressed to Meiqi Li, Peng Xi or Dayong Jin.

**Peer review information** *Nature Communications* thanks David Baddeley and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

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

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