DeepCristae, a CNN for the restoration of mitochondria cristae in live microscopy images

Salomé Papereux[1,2,&], Ludovic Leconte[1,2,&], Cesar Augusto Valades-Cruz[1,2,&], Tianyan Liu[3], Julien Dumont[4], Zhixing Chen[3], Jean Salamero[1,2], Charles Kervrann[1,2], Anaïs Badoual[1,2,*]

1 SERPICO Project Team, Centre Inria de l'Université de Rennes, F-35042 Rennes, France

2 SERPICO Project Team, UMR144 CNRS Institut Curie, PSL Research University, F-75005, Paris, France

3 College of Future Technology, Institute of Molecular Medicine, National Biomedical Imaging Center, Beijing Key Laboratory of Cardiometabolic Molecular Medicine, Peking-Tsinghua Center for Life Science, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

4 CIRB Microscopy facility, Collège de France, UMR 7241 CNRS, Inserm U1050, Paris, 75005, France

[&] these authors contributed equally to the work

* corresponding authors contact: anais.badoual@inria.fr

Abstract

Mitochondria play an essential role in the life cycle of eukaryotic cells. However, we still don't know how their ultrastructure, like the cristae of the inner membrane, dynamically evolves to regulate these fundamental functions, in response to external conditions or during interaction with other cell components. Although high-resolution fluorescent microscopy coupled with recently developed innovative probes can reveal this structural organization, their long-term, fast and live 3D imaging remains challenging. To address this problem, we have developed a convolutional neural network (CNN), called DeepCristae, to restore mitochondrial cristae in low spatial resolution microscopy images. Our CNN is trained from 2D STED images using a novel loss specifically designed for cristae restoration. Random sampling centered on mitochondrial areas was also developed to improve training efficiency. Quantitative assessments were carried out using metrics we derived to give a meaningful measure of cristae restoration. Depending on the conditions of use indicated, DeepCristae works well on broad microscopy modalities (STED, Live-SR, AiryScan and LLSM). It is ultimately applied in the context of mitochondrial network dynamics during interaction with endo/lysosomes membranes.

Keywords: Convolution neural network, image restoration, mitochondria cristae, fluorescence microscopy, super resolution, live imaging

Introduction

The study of certain pathologies has shown the importance of mitochondria, which above all, ensure ATP production within cells and are central in many biological functions (e.g., metabolic pathways, ion homeostasis, apoptosis, autophagy, epigenetics…)[1,2]. Mitochondrial energetic adaptations to environmental constraints encompass a plethora of processes that maintain cell survival. An alteration of these processes generally leads to serious diseases such as cancer, neurodegenerative and cardiovascular disorders[3]. Although much attention has been paid to the role of mitochondria, the precise niche the organelle plays in cell life and death still remains unclear. The lack of in-depth knowledge about the ultrastructural evolution of mitochondria in live cells, under normal and stressful conditions, might be one of the blind spots. In particular, the cristae formed by the inner membrane of mitochondria that concentrate ATP production in a defined area, their dynamic behavior, sublocation or density have been poorly related to the various functionalities or dynamic processes (e.g., fusion, fission) that mitochondria undergo. The challenge we address, lies in imaging mitochondria cristae, which measure between 30 and 50 nm wide[4], at a high spatial and temporal resolution so that their structural dynamics and interactions can be accurately studied over time for several dozens of milliseconds to a few seconds. However, this is starting to be possible with the recent development of high-resolution imaging approaches[5].

Stimulated emission depletion (STED) microscopy, which allows for sub-diffraction resolution (xy: 30-50 nm), is one of the very few techniques[6,7] able to decipher dynamics of mitochondria cristae in live cells[4]. However, their observation in 3D and in fast time is limited by the acquisition frame rate capacity (1 plane ≈1 to 10 s). In addition, depletion STED, which is the principle that achieves nanoscopic resolution, induces local heat by high illumination intensity[8] to which mitochondria are known to be particularly sensitive[9,10]. This can affect their overall physiology and potentially lead to apoptosis and mitophagy. A number of new fluorescent probes that are more photostable with less saturation intensity and that allow cristae decoration, have been developed the very last years[7,9,11,12]. Yet, the application of a dark recovery step (≈30 s) after STED imaging is still necessary, again at the expense of temporal resolution. This could be improved by applying a partial STED depletion protocol, leading to an intermediate quality resolution (xy≈100 nm)[13], but insufficient to spatially resolve mitochondria cristae and not solving the frame rate limitation (4-5 s in average).

In this context, one solution to study the dynamics of mitochondria cristae is to collect as much temporal information with minimal phototoxicity using an appropriate microscope, and then restore the spatial dimension using computational methods. Indeed, the development of image restoration algorithms has become increasingly popular in recent years with the need for nanoscale analysis[14–22]. At the heart of fluorescence microscopy have been actively developed denoising algorithms[23–28], dedicated to images corrupted by a mixed Poisson-Gaussian noise, as well as deconvolution algorithms[29–31], designed to remove the blur induced by the limited aperture of the microscope objective. Some methods combine the two approaches[32]. However, these conventional restoration methods usually rely on general assumptions, such as the nature and level of noise and spatial regularity, which hampers their effectiveness on the diversity of structures and level of degradation in microscopy images. Over the years, the literature on image restoration has evolved considerably due to deep learning and the rapid growth of convolutional neural networks (CNNs). These methods have the advantage of making assumptions based on image content, resulting in state-of-the-art performance in

denoising[33,34] and deblurring[14,19,22,35–38] fluorescence microscopy images. However, these methods have two major drawbacks. First, these CNNs often require a training step based on a large ground truth dataset that is generally not available in microscopy. Second, they focus on restoring the entire image, while sometimes little information is worth restoring within it, especially in the dark background. This is the case with mitochondria cristae, which have a sparse number of pixels in the image compared to the background. Therefore, CNNs that have been previously applied to mitochondria microscopy images[21,22,38,39] provide good global restoration of the background and mitochondria but fail to accurately restore fine details as cristae, especially in very low spatial resolution images. To circumvent this, new conventional methods have been proposed to enhance resolution and suppress artifacts in high-resolution techniques, including Hessian-SIM[17]. However, as explained above, the denoising results are limited when dealing with low signal-to-noise ratio images and Hessian deconvolution assumes that the unknown image is smooth in some sense and sparse. A hybrid solution has been proposed in TDV-SIM[40], which combines the strengths of conventional physical model-based algorithms with deep learning-based algorithms. Another hybrid solution, rdLSIM[21], incorporates the deterministic physical model of specific microscopy into network training and inference. Nevertheless, the effectiveness of these methods, along with conventional restoration algorithms, relies on the careful selection of optimal parameters or on prior knowledge of illumination patterns, respectively.

Instead of developing an additional generic image restoration method that may not satisfactorily enhance certain sparse but informative pixels in the image, we present DeepCristae, a CNN specifically developed to restore mitochondrial cristae in low spatial resolution microscopy images. DeepCristae was applied to several microscopy modalities and different biological scenarios capturing live mitochondria at high speed with low illumination and thus low phototoxicity. DeepCristae allows long-term/fast dynamic observation of cristae behavior and organization. The main challenge was to handle the low number of cristae pixels compared to the background in the acquired images. Therefore, the main contributions of this work are 1) the design of a new training loss dedicated to the restoration of specific pixels of interest, 2) a random patch sampling focusing on areas of mitochondria, and 3) the building of metrics for objective assessment of cristae restoration.

Results

Overview of DeepCristae

DeepCristae aims to restore mitochondrial cristae in intermediate to low spatial resolution microscopy images (Fig. 1). DeepCristae mainly consists of a U-Net trained on a dedicated dataset built from real high-resolution 2D STED images (Fig.1 and Supplementary Table 1) and using a novel training loss we specifically designed for cristae restoration (Methods, Eq. (1)). Although the term is not fully appropriate, for simplification we refer to this dataset as "synthetic" D_{synt}. A pipeline for random patch sampling focusing on regions of mitochondria in the acquired data was also developed (Supplementary Fig. 1). DeepCristae image restoration network was implemented in Python (TensorFlow version 2.11) and is freely available as an open-source software (see code availability). DeepCristae is also integrated into BioImageIT[41], an open-source platform with existing software for microscopy.

DeepCristae quantitatively outperforms state-of-the-art algorithms on the synthetic dataset D_{synt}

Our method was quantitatively compared to existing both conventional and deep learning algorithms for image restoration. It includes denoising methods (ND-SAFIR[23], 2D median filter and Noise2Void[33]), deconvolution algorithms (Richardson-Lucy[29,30] and Wiener[31]), and approaches combining both, such as SPITFIR(e)[32], CARE[19], RCAN[37] and SRResNet[35] (details are in Supplementary Note 2.2.2). All methods were applied to 9 test images of D_{synt} (Supplementary Note 1.2.1). Note that in this dataset, the ground truth images were obtained by deconvolving the high-resolution 2D STED images with a Richardson-Lucy (RL) algorithm to enhance mitochondria cristae. Iterative RL algorithm is suitable for signal-dependent noise removal. Nevertheless, after several iterations it may create unwanted artifacts such as a "night sky" pattern in the image. To generate ground truth images from the input 2D STED images, restored intermediate images with little “sky night” pattern, less blurred and noisy than the input images, were carefully selected. All deep learning methods were trained from the same patches extracted from D_{synt}.
First, to evaluate the performance of the different methods, we used current metrics, namely NRMSE, PSNR and SSIM (Supplementary Note 2.1). However, these measures are relevant to the image as a whole, but insufficient in the context of mitochondrial cristae restoration. Indeed, the images contain only few pixels of cristae and thus have too little impact in those metrics unlike the many background pixels. To overcome this issue, we encouraged the evaluation metrics to focus exclusively on mitochondria pixels. To solve this problem, evaluation metrics are focused exclusively on mitochondrial pixels. We call these mitochondrial metrics $NRMSE_{mito}$, $PSNR_{mito}$ and $SSIM_{mito}$. To go one step beyond and accurately assess cristae restoration, we also introduced the cristae metrics $NRMSE_{cristae}$, $PSNR_{cristae}$ and $SSIM_{cristae}$. These metrics are computed over mitochondria cristae pixels only, obtained from manual annotations (Supplementary Note 2.1). Each competing algorithm was evaluated over nine test images, for the nine aforementioned metrics (Fig. 2b). For all measurements focusing on mitochondria or cristae, DeepCristae ranks first, and is either first or second otherwise. Conventional methods behave worse than deep learning approaches, CARE appearing DeepCristae's most competitive method. However, as confirmed by the values of the metrics $NRMSE_{cristae}$, $PSNR_{cristae}$ and $SSIM_{cristae}$, CARE restores mitochondria cristae with less sharpness, especially for mitochondria with low contrast or in a noisy background (Fig. 2a, CARE white arrows). It should be noted that if we had considered conventional NRMSE, PSNR and SSIM measurements, we would not have been able to demonstrate the benefits of restoring more informative pixels. In terms of visual assessment, SRResNet and RCAN amplify the background noise, resulting in less accurate restoration of cristae and unrealistic reconstructed structures in the background or in mitochondria (Fig. 2a). DeepCristae removes noise background while restoring most of the cristae details.

Second, a test image of D_{synt} was restored with DeepCristae (Fig. 3a). We selected four regions of interest (ROIs) to better appreciate the restoration results (Fig. 3b-e). For each ROI, a comparison of normalized intensity profiles between the input image, DeepCristae restored image and the ground truth image is performed. It shows that DeepCristae restores spatial information by revealing mitochondria cristae while improving signal to noise ratio. The line profiles of DeepCristae are sometimes smoother than the ones of the ground truth, especially in areas where the image background is noisier (Fig. 3d, e). This is due to a residual “night sky” effect in the ground truth that does not appear with DeepCristae. This surprising result

can be explained in two ways. First, we reinject information from the input into the network at the end and second, our new Similarity Component Prioritization (*SCoP*) loss (see Methods) especially focuses on the mitochondria and cristae pixels where there are the fewest artifacts with RL.

Robustness and stability of DeepCristae.

We have shown that DeepCristae performs well on 2D STED images and outperforms state-of-the-art algorithms. However, it is important to verify the reliability of DeepCristae more widely. DeepCristae has been trained on a dedicated dataset acquired with specific microscope settings and mitochondria properties (e.g., fluorescence markers, width in pixels of the mitochondria in the images). Any change in these settings is expected to alter the quality of the restoration results. To evaluate the influence of changes in these parameters on the results, three experiments were performed. Our model was trained from the training images of D_{synt} depicting mitochondria of width 15.64 ± 4.04 pixels on average. We started by studying the quality of the predictions as a function of the mitochondria width in pixels in the input images. To that end, the 9 test images of D_{synt} were rescaled 11 times in order to contain mitochondria of specific widths (in pixels) on average. It thus results on 11 test datasets on which our trained DeepCristae model was applied (Supplementary Fig. 2a, b) and the metrics (Supplementary Note 2.1) were computed. The evolution of the metrics as a function of average mitochondrial width shows that the closer you get to the training parameters (i.e. an average width of 15.64 pixels), the better the quality of the restorations. In fact, if the mitochondria are too small, few cristae are restored and the mitochondria are thin. On the contrary, if the size is too large, DeepCristae tends to create artifacts looking like cristae patterns. Next, our model was trained on images obtained with specific parameters that mimic microscope settings. Real images are assumed to be corrupted by mixed Poisson-Gaussian noise (with standard deviation $\sigma_{noise} = 4$) and the point spread function of the microscope is approximated by an isotropic Gaussian function of standard deviation $\sigma_{blur} = 3.25$ pixels. We performed two experiments, similar to the one described above, to examine the assessment of DeepCristae image restoration as a function of σ_{noise} and σ_{blur} (Supplementary Fig. 3). For both experiments, visual results show that the quality of the restoration decreases as σ_{noise} and σ_{blur} increase (Supplementary Fig. 3b, d). The higher the σ_{noise} or σ_{blur} values are, the blurriest the mitochondria's boundaries and their cristae. This is also confirmed by the evolution of the metrics as a function of σ_{noise} (Supplementary Fig. 3a) and by the evolution of $SSIM_{mito}$ and $SSIM_{cristae}$ as a function of σ_{blur} (Supplementary Fig. 3c). Surprisingly, the evolution of the PSNR and NRMSE as a function of σ_{blur} have a bell-shape with a maximum and a minimum, respectively, for values of σ_{blur} close to 3.25 pixels. We thus recommend using DeepCristae on microscopy images with blur and noise levels at worst equal to our training conditions ($\sigma_{noise} = 4$ and $\sigma_{blur} = 3.25$ pixels). Beyond this, the quality of the restoration can drastically decrease, especially for blur, and for too high values of these parameters DeepCristae could reveal hallucinated cristae shapes.

Moreover, it is important to guarantee that under well-controlled conditions of use (mitochondria width, image blur and noise), DeepCristae is hallucination-free and stable. By stable we mean that different training leads to consistent predictions. To demonstrate that these requirements have been met, we performed three experiments. First, we trained 10 DeepCristae neural networks with different training data, each generated with our patch generation method applied to the 24 training images of D_{synt} (Fig. 4g-l). For each training, the

resulting model was applied to the 9 test images of D_{synt} and the aforementioned metrics were computed. The average metrics obtained over the 10 trainings are close to the ones obtained with our model and the standard deviations are very low, showing consistency between predictions (Fig. 4l). By visually analyzing the predictions, the color map of the standard deviation (Fig. 4g) as well as looking at normalized intensity line profiles along mitochondria (Fig. 4h-k), we observe that the 10 trainings agree overall on the presence or absence of cristae but diverge in their intensity and their precise boundaries. In this experiment, all networks were initialized with the same weights, confirming that our patch sampling method is robust and leads to homogeneous learning. A second similar experiment was carried out. Ten trainings were performed from one dataset but with 10 different weights' initializations (Supplementary Fig. 2c-h). This experiment indicates that the same dataset leads to homogeneous learning, meaning that the randomness of initialization does not play a key role in the learning process. Finally, to qualitatively assess the performance of DeepCristae on real data, we acquired 5 pairs of real 2D STED images. Each pair contains one low-resolution (LR) and one high-resolution (HR) STED image, acquired as quickly as possible (~30 s), to minimize the displacements and deformations of mitochondria between the two acquisitions (Supplementary Note 1.2.1). The HR STED images were deconvolved using the RL algorithm to enhance mitochondria cristae and are considered as "ground truths" (GTs). The LR STED images were resized to have an average mitochondrial width of 15.64 pixels (391 nm), in line with the conclusions drawn above, and were then given as input to DeepCristae. The obtained predictions were qualitatively compared to the GTs to control their consistency. Four ROIs, from two of the five pairs of real STED images and corresponding predictions, were selected in regions where small mitochondrial displacements were observed to better appreciate the restoration (Fig. 4a-f). For each ROI, a comparison of normalized intensity profiles between the input LR image, DeepCristae restored image and the ground truth image is performed. A consistency between the cristae restored by DeepCristae and the ones present in the ground truths is observed overall, and no meaningful "hallucination" is observed.

DeepCristae enables to restore 2D+time STED images

While 2D STED nanoscopy enables to resolve mitochondria cristae and was here helpful to develop DeepCristae, live STED acquisition encompasses a number of hurdles. It includes relatively long-time frames between images, even when a photostable probe was used (Fig. 5a, b), limiting the temporal overview of the mitochondrial dynamics in the same plane. Moreover, STED imaging may rapidly induce photo-bleaching, which makes ultrastructural details progressively dimmed. More problematic, repeated STED imaging rapidly induces morphological deterioration of live mitochondria, illustrated by their swelling in the latest time points (Fig. 5a, b). This swelling effect was quantified here for HR STED by measuring the lateral widths of 7 distinct mitochondria over the 10 time points (Fig. 5d) in the image series (Fig. 5a, b). The swelling appears between the 5th and the 7th frames.

In order to improve the frame rate of 2D+time STED imaging while limiting the photodamages on mitochondria, one may adjust the STED imaging protocols (Supplementary Note 1). Accordingly, we built another live STED dataset, first to give an indication on how long, with STED acquisition in lower depletion conditions, we can image before the mitochondria are damaged and second, to control the efficiency of DeepCristae restoration over time. In what follows, these low-resolution (LR) STED images were referred to as Fast STED images. This goes at the depend of *xy* resolution (Fig. 5c, left bottom triangles in image series) for both the

lateral width of mitochondria and the cristae width (Fig. 5d, Fast STED and Fig. 5e, RAW, respectively). Applying DeepCristae restoration on these latest series clearly revealed cristae morphology (Fig. 5c right top triangle in image series). As expected, LR (or Fast) 2D+time STED images show little changes in mitochondria lateral widths in time, in contrast to HR STED (early and late time points in Fig. 5d) but a degraded resolution in the cristae widths (from a mean (μ) of ~90nm in HR STED to ~120 nm in Fast STED with σ = ± 47 nm). Applying DeepCristae allows recovery of a resolution lower than 100 nm and drastically reduces the variability of the measurement. The mean crista-to-crista distance, measured as peak-to-peak intervals (Fig. 5f), widely depends on the cristae density along the mitochondria network. Here, in RPE1 cell, it varies from 50 nm to more than 500 nm in early time points in HR STED (μ =319 nm, σ = ± 247 nm) while the heterogeneity increases in late time points (from 130 nm up to 1.6 µm), consistently with the observable swelling of the mitochondria. In Fast STED, the cristae intervals measurements were non-significant. However, after DeepCristae restoration the mean crista-to-crista distance was estimated at ~142 nm (σ = ± 46 nm) (Fig. 5f, g). Differences in these crista-to-crista measurements with similar studies on HeLa or Cos7 cells for instance, will be further discussed. DeepCristae restored the individual cristae at 81 nm of resolution (σ = ± 9 nm) at FWHM (Full Width at Half Maximum), as compared to the approximately 50 nm obtained in other studies[4]. DeepCristae provides a useful way to improve live STED nanoscopy by improving the resolution and decreasing the frame rates (3 to 6s versus 13s), yet with no observable photodamage as illustrated here by measuring the swelling of mitochondria.

DeepCristae restore 3D+time images of mitochondria cristae by using intermediate high-resolution and diffraction limited microscopy

STED nanoscopy is not the only high-resolution microscopy adapted to resolve internal mitochondria ultrastructures in live cells. Indeed, a number of works using adaptation of SIM approaches have been published over the last few years[16,42], some combined with conventional deep learning methods[20,22,39]. Yet, the best compromise between Fast and 3D imaging still remains an issue. We next investigated the performance of DeepCristae prediction on intermediate HR microscopies chosen for their 3D optical slicing performance. Spinning disk confocal equipped with a Live-SR module (or SDSRM for Spinning Disk Super Resolution Microscopy) is one of those well-disseminated systems equivalent to SIM. It improves the *xy* resolution by a factor of ~2 (~120-130 nm at 488 nm, ~140 nm at 561 nm)[43] while giving access to the depth (z-axis) of the sample and the live imaging of mitochondria (time t) without severe photo-bleaching and phototoxicity. The use of the Live-SR is therefore motivated here by both the study of these four dimensions and the ability of our model to correctly perform cristae reconstruction via multiple microscope imaging modalities. DeepCristae efficiently revealed cristae organization in single 2D Live-SR images acquired within 30 ms (Fig. 6a, upper images) and thus in 3D (Fig. 6a, lower images, MIP on 14 planes, with a stack time ~800 ms), giving access to the overall mitochondria network in the live cell at a fast rate. In this respect, it outperforms HR STED imaging and even Fast (LR) 2D-STED imaging after DeepCristae restoration (compared to Fig. 5c). Cristae width comparative estimation (Fig. 6b) shows the improvement in resolution obtained after DeepCristae restoration on single plane Live-SR images (Raw =149 ± 64 nm when measurable; DeepCristae =87 ± 11 nm). These results are close to the expected widths of circumvoluted cristae tubules (50 to 100 nm) obtained by other methods derived from SIM imaging[22]. We then tested DeepCristae restoration on LLSM[44] imaging which is not an HR microscopy by

itself (at least in the dithered mode) but gives the best compromise in terms of fast 3D acquisition with minimal photon dose illumination and consequently low photo-damage of the mitochondria over time. Surprisingly, although with intrinsic limited and non-isotropic resolutions (PSF xy=300 nm and z=600-700 nm in our system) and a particular geometrical acquisition mode, cristae were however detectable in some mitochondria after realignment and a Richardson Lucy (RL) deconvolution. The resulting images were here considered as "Raw" data (Fig. 6c, 2D upper panel, left image). Applying DeepCristae to them (Fig 6c, 2D upper panel, right image and composite zoomed area for comparison) improves the cristae resolution (Raw = 339 ± 248 nm, when measurable; DeepCristae= 94 ± 15 nm) and strongly reduces the variance of paired measurements (Fig. 6d). One of the obvious advantages of LLSM over confocal imaging is to allow continuity between single image planes over large stacks coupled to an extended depth of focus, as illustrated here by the 3D rendering as an oblique projection (Fig. 6c**,** 3D). Moreover, LLSM is particularly adapted to long range/high frequency imaging on whole living cells, which, coupled to low photon dose illumination, makes it one of the best imaging systems, if not the best, for the highly light-sensitive organelles that are the mitochondria. Applying DeepCristae adds information on cristae ultrastructural organization in the whole mitochondria network of the cell. Finally, Fast 3D Live-SR and LLSM time series (Fig 6e, f) were treated for DeepCristae restoration. Cristae ultrastructural features can be observed, while the mitochondrion network undergoes well known dynamic modifications such as fusion or fission processes (Fig. 6e, f, panels of composite zoomed area in both time series; left "RAW " and right "DeepCristae"; Movies S1 and S2). Images are of better quality after restoration of Live-SR compared to LLSM images. However, it should be noted the gain in acquisition parameters for the latter in these experiments, with 75 slices per stack and a double channel stack time=1.3 s versus 14 planes per stack and double channel stack time= 3.9 s for Live-SR. DeepCristae restoration was also tested with an AiryScan 5 LSM 980. It provided similar improvements, although for a 15 planes stack time of about 30s and with more artifacts appearing after DeepCristae, the nature of which most probably lies in the way the reconstruction of the AiryScan images was carried out from the values determined automatically by the commercial software (Supplementary Fig. 5).

DeepCristae restoration allows to decipher mitochondria cristae morphodynamics during inter organelles interactions

The most documented membrane-membrane interactions involving mitochondria are the endoplasmic reticulum (ER)–mitochondria contacts, whose functions have been continuously expanded since the 1990s[45,46]. In addition to the ER, mitochondria contact vacuoles/lysosomes, peroxisomes, lipid droplets, endosomes, the Golgi, the plasma membrane (PM) and melanosomes[47]. The number of these interactions as well as their duration drastically vary from one type to the other, as they depend on the respective membrane surface of the specific organelles within the cell and their contact time[48]. Their detection may thus require fast and/or long-range 3D imaging. As already mentioned, even high-resolution approaches which are well adapted to decipher ultrastructural features of mitochondria such as cristae, generally fail to capture their dynamic evolution in the 3D space of the whole cell at multiple time scales. This can be critical, if one wants to study inter-organelle membrane interactions and their effects. We next initiate the investigation of endosome/lysosome-mitochondria dynamic interactions by addressing specifically the ultrastructural behavior of the cristae during these contacts. This was done by imaging multiple 3D+time double fluorescence series in Live-SR (represented as a single stack MIP in Fig. 7a,

top left) or LLSM (represented as a single stack MIP in Fig. 7c**,** top left), where the membranes of the endo-lysosomal pathway were continuously labeled (Supplementary Note 1.1). DeepCristae restoration was applied on both MIP datasets (respectively, Fig. 7 a, c, bottom left). A number of mitochondria dynamic events correlated with endosomal structure behaviors were captured. Only a few of them are here extracted as thumbnail time series (Fig. 7b, d) of zoomed area (colored insets in Fig. 7a, c) from the Live-SR and LLSM acquisitions, respectively. Among others, the formation of endo-lysosomes contacts sites with mitochondria (Fig. 7b, blue and green), positioning of endo-lysosomes relative to fission sites of mitochondria (Fig. 7b, blue and red), very long confinement of endo-lysosomes within the mitochondria network (Fig. 7b, orange) and image series of endo-lysosomes appearing to pull a small mitochondria from one to another elongated tubules of mitochondria (Fig. 7b, red). DeepCristae restoration on the space-time localization of these events can also be evaluated dynamically (Movie S3). Similar events are followed with LLSM, such as the fission of mitochondria at a contact site with an endo/lysosome vesicle (Fig. 7d, orange) and long confinement of an endo/lysosome vesicle within the mitochondria network (Fig. 7d, green). The main advantage of the LLS modality (fast frame rate, low photon illumination of the sample coupled to whole cell 3D acquisition) is the improvement of the time resolution of the data series (or long-range acquisition). Consequently, fast events involving endo-lysosome contacts with mitochondria are easier to capture and these dynamics are precisely deciphered. For instance, one may extract first (Fig. 7c, blue inset), probably a fusion process (Fig. 7d, from time point 105 to time point 119; $\Delta t = 13\ s$), and second a fission process (Fig. 7d, from time point 192 to time point 198; $\Delta t = 6\ s$; Movie S4). At each time point, DeepCristae-restored mitochondria and denoised/deconvoluted endo-lysosomes double-labeled images (Supplementary Note 1) are paired to non-treated images (right and left panel, respectively, of thumbnails time series in Fig. 7b, d). While cristae resolution in LLSM does not reach that obtained with Live-SR, DeepCristae restoration brings values closer together (Fig. 6b, d). In all situations and for both intermediate HR (Live-SR) and diffracted limited (LLSM) imaging modalities, DeepCristae restoration provides ultrastructural information on the positioning, density, and dynamics of mitochondria cristae. We then wanted to quantitatively assess how the dynamic architecture of the mitochondria internal membrane during endo/lysosomes-mitochondria interaction could be revealed with DeepCristae. We focused on the fission process. To do so, we first selected 21 distinct 3D+time image series from the Live-SR datasets, in which mitochondria fission was monitored. Intensity line plots were measured along mitochondria on some time points framing the fission event. This was done on both DeepCristae-restored and unrestored individual time points in a “blind” manner, meaning without looking in the second channel depicting the location of endo-lysosomes (Supplementary Fig. 6a). Measurements of “peak to peak” intervals between cristae, were only possible in the DeepCristae restored images and show an increased density after fission occurs (Supplementary Fig. 6b, dark circles). Interestingly, by overlaying the second channel in a second step, 60% (13 over 20) of these selected time series showed proximity if not direct contact at the site where mitochondria fission is observed (Supplementary Fig. 6b, red circles; 6c for statistics). While still preliminary and not deciphering the exact nature of the endosomal compartments involved (i.e., PMDR labels the overall endo-lysosomal pathway), this illustrates how DeepCristae would represent an asset to quantitatively study the dynamic architecture of the mitochondria internal membrane during diverse dynamic processes or in particular physiological or constrained conditions.

Discussion

Mitochondrial membrane architecture is essential for the many functions of mitochondria. In particular, mitochondria cristae are the main site of energy production and are dynamic ultrastructures remodeling upon various cellular stimuli and natural processes (apoptosis[1]; aging[49]). Therefore, understanding the structure and dynamics of cristae is vital for comprehending mitochondrial function and its implications in cellular physiology and diseases. High-resolution microscopy coupled with robust mitochondrial probes[7,9] are key recent developments that started to reveal the fine details of mitochondrial cristae structure and organization, overcoming the limitations of conventional microscopy. However, imaging at high spatial and temporal resolution remains a challenge.

DeepCristae exploits the power of deep learning to reveal cristae in images taken with low photon illumination, enabling clearer visualization and analysis of mitochondria cristae in living cells without interfering with the natural behavior of mitochondria. While it has been trained on a dedicated dataset that was created from real high-resolution 2D STED images, we have shown that it operates for a wide range of optical resolutions, from diffraction-limited to intermediate high-resolution microscopy, providing researchers with a powerful tool to study cristae dynamics without compromising their structural integrity or functionality.

While there are other deep learning approaches available for revealing cristae ultrastructure[21,22,38,39], DeepCristae offers unprecedented advantages. First, thanks to a well - defined training loss dedicated to the restoration of mitochondria signals, it outperforms state-of-the-art methods. Secondly, it not only makes it possible to visualize and restore cristae dynamics in 2D STED nanoscopy with minimal illumination and without damaging mitochondria but more importantly, it extends these capabilities to other high-resolution imaging techniques such as Live-SR, SIM and AiryScan, more suited to such 3D dynamics. Finally, DeepCristae can be applied to advanced microscopy techniques such as LLSM, enabling fast and long-duration 3D+time acquisitions within the diffraction-limited range. This versatility makes DeepCristae a unique and valuable solution for studying cristae dynamics across a range of spatial and temporal scales.

Overall, our results show that fluorescence microscopy combined with DeepCristae enables long-term/fast dynamic observation of cristae behavior and organization with high quality. To illustrate the contribution of our approach to biological phenomena that are likely to involve the functional structure of mitochondria, we have chosen to focus on inter-organelle interactions and their consequences. While mitochondria-associated ER membranes, the biochemical composition of the contact sites and diverse physiological and disease-related functions have been extensively studied over the decades[50,51], it is increasingly recognized that other organelle contacts have a vital role in diverse cellular functions[52]. More recently, there has been growing interest in quantifying other membrane interactions with mitochondria and their cell distribution in space and time[48], in particular within the endo-lysosomal pathway and their contribution to the fission/fusion process of the mitochondria network[53]. Here, while confirming the coincidence of contacts between the endo-lysosomal membrane and mitochondria, we enlightened the change of cristae density during fission (Supplementary Fig. 6). This density as well as complex cristae arrangements depends on cell types and metabolic activities[4,54], not talking of obvious modifications induced by environmental conditions. Until now, to provide a dynamic view of individuals and groups of cristae required 3D nanoscopy or linear SIM[16],

which are not always compatible with the time frame required to capture the event of interest. In this respect, DeepCristae might be an asset to compare the cristae dynamics in different cell types and in these various conditions.

However, as with any image restoration method, scientists may be concerned by the robustness of DeepCristae to accurately restore mitochondria cristae and not hallucinate them. This is why we investigated the stability and limits of our method. First, we consider the RL algorithm in our proof-of-concept as it may be considered as the baseline algorithm for image deconvolution in fluorescence microscopy. The behavior of this algorithm is well established, including the "night sky" artifacts when iterated. The GTs shown in the manuscript may be improved if we consider more sophisticated deconvolution algorithms. We also worked out different conditions of use to be respected to guarantee good quality and truthfulness of the results. It is important to feed DeepCristae with images containing mitochondria whose average width in pixels is close to the one seen during the training. Concerning the microscope settings, it is better to ensure that the level of noise and blurring in the input images is equivalent to or better than the one present in the training data (which was quite high in our training). Under these conditions of use, across all our experiments and through different microscopy modalities, no hallucination was observed: a consistency between line profiles along mitochondria between raw and restored data was always observed.

Like cytoskeletal elements, the mitochondrial ultrastructure is a key element for comparing the performance of new super-resolution microscopy techniques. In terms of applications, DeepCristae makes it possible to track the evolution of mitochondrial cristae morphology over time, during interactions with other membrane components of the cell, or under extracellular conditions that mimic various pathological or stress situations.

Methods

In this section, we present the main features of DeepCristae. We first present the dataset we created from real 2D STED images to train the network. Then, we overview our network architecture and present the novel learning loss function, which prioritizes the restoration of specific pixels. We finally detail the image patch-sampling steps, which is a crucial stage in the model pipeline.

Generation of the 2D STED dataset - D_{synt}**.** As mitochondria are living organelles, mostly organized as a quite fast-moving network in RPE1 cells (Supplementary Note 1.1), the acquisition of a pair of high and low-resolution images at the exact same time point is impossible. To train and quantitatively validate DeepCristae, we thus created a dataset, called D_{synt}, from 33 acquired high resolution (HR) 2D STED images (25 x 25 nm) that we denote I_{HR} (Fig. 1 a-d). More information on the acquisition of the images I_{HR} are available in Supplementary Note 1.2.1.

First, we degraded the images I_{HR} to obtain low resolution (LR) images of mitochondria, denoted $I_{synt_{LR}}$, that will serve as input to the neural network. To that end, we first applied a Gaussian filter of standard deviation $\sigma_{blur} = 3.25$ pixels to the images I_{HR} in order to approximate the blurring effect due to the point spread function of the microscope. Then, we added a Poisson-Gaussian noise ($\sigma_{noise} = 4.0\ pixels$), consistently with noises observed in real STED images. The parameters σ_{blur} and σ_{noise} were set to create pertinent input data that mimic real LR STED images (Supplementary Note 1.2.1).

Second, as our goal is to learn how to restore LR images to better reveal the mitochondria cristae, we paired the LR STED images $I_{synt_{LR}}$with their restored counterpart $I_{synt_{GT}}$. These ground truths are obtained by deconvolving the real HR images I_{HR} with the Richardson-Lucy algorithm[29,30], enhancing the mitochondria cristae. Other deconvolution algorithms were tried, such as SPITFIR(e)[32] or Wiener[31], but the results obtained after training were not as good.

Finally, the dataset D_{synt} is made of 24 training images and 9 test images. To further increase the size of the training set, data augmentation and patch sampling (later described in Methods) are performed on the pair of LR images $I_{synt_{LR}}$ and ground truths $I_{synt_{GT}}$. The dataset is first augmented by applying three different rotations to the images (90°, 180° and 270°). Then, a shrink transform, and horizontal and vertical flips are successively applied to 25% of the augmented dataset, randomly selected. The final training set is made of 1824 patches of size 128×128 pixels, whose 20% are used for the validation set (summary in Supplementary Table 1).

Network architecture. We used the network proposed by Weigert et al.[19] as the backbone of the CNN architecture, itself built upon the U-Net[55]. It has a contracting path and an expansive path, each one consisting of 3 sequential downsampling and upsampling blocks, respectively. Each block of the first path is skip-connected to the associated one of the expansive paths. The contracting path consists of two successive 3×3 convolutions, each followed by a Rectified Linear Unit (ReLU), and a 2×2 max pooling operation with stride 2 for downsampling. Every depth in the expansive path consists of a 2×2 up-sampling of the feature map, concatenated with the corresponding feature map from the contracting path, followed by two 3×3 convolutions with a ReLU activation function. At the final layer, one 1×1 convolution is used. The output results from an additive assembly between the input of the neural network and the last layer's output. The network (Fig. 1f) outputs the same size restored images.

Design of the training loss. We present our new loss, the Similarity Component Prioritization (*SCoP*) loss, that has been designed to better restore mitochondria cristae. Most losses and metrics used to train networks or to evaluate the quality of restorations compute the score on the whole image, giving the same weight to any pixel. For example, the *MAE* computes the mean absolute error between the prediction and the target image, while the *SSIM*, despite not basing the calculation on pixels-to-pixels difference, computes the similarity among all the pixels of both images. Instead, our purpose is to focus on informative pixels corresponding to target structures in images. Indeed, the dark and noisy background occupies most of fluorescence images, which alters the learning. To overcome this issue, we introduce *SCoP*, a novel loss which adaptively prioritizes the restoration of mitochondria pixels.

Our *SCoP* loss is built upon the structural dissimilarity (*DSSIM*) measure. Consider (i, j) the spatial coordinates of a given pixel and a patch of size (N, N). The loss formula between a target image y and its prediction $\hat{y}$ is given by

$$SCoP(y, \hat{y}) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\frac{1 - SSIM_{y,\hat{y}}^{map}(i, j)}{2} \right)^{\gamma_{i,j}}, \quad (1)$$

Where $SSIM_{y,\hat{y}}^{map}$ is the map of the local structural similarity (*SSIM*) values for corresponding pixels between the images y and $\hat{y}$. Each *SSIM* value ranges in [-1,1], where -1 (1,

respectively) testifies of a bad (very good similarity, respectively) between $y(i,j)$ and $\hat{y}(i,j)$. The parameter $\gamma_{i,j}$ prioritizes the restoration of specific regions of interest. In our case, we chose $\gamma_{i,j} = 1$ if the pixel (i,j) belongs to a mitochondrion, 4 otherwise. In this way, we encourage the network to focus on restoring mitochondria pixels and reduce the influence of a poorly restored background on the loss. Determining whether a pixel belongs to the background or to a mitochondrion can be performed automatically (using our method described in Methods "Image patch sampling for the training step - Thresholding") or manually by using any binary segmentation provided by the user.

Data normalization. Our training images of D_{synt} have different ranges of intensity values. To homogenize them, we normalized the input data and their corresponding ground truth to a common distribution of intensity values with the percentile normalizer. This normalization also has the advantage to exclude outliers, which are very frequent in microscopy imaging due to noise and luminance. The percentile normalization of an image I is defined as

$$I_{norm} = \frac{I - perc(I,\, p_{low})}{perc(I,\, p_{high}) - perc(I, p_{low})}, \tag{2}$$

where $perc(I,p)$ is the *p*-th percentile of I. We used $p_{low} = 2$ and $p_{high} = 99.8$. This step is also performed during the inference step on any input data.

Image patch sampling for the training step. Our model is trained on the dataset D_{synt} containing 24 images (96 after data augmentation) of different sizes. In order to homogenize and increase the training dataset, we performed patch sampling. We sampled each input training image $I \in R^{W \times L}$, defined over the grid Ω of size $W \times L$, within $N_I = \left\lceil \frac{W}{128} \right\rceil * \left\lceil \frac{L}{128} \right\rceil$ patches of size 128×128. As our images contain more background pixels than mitochondria pixels, grid or simple random patch sampling will end in too many empty patches. This can degrade the training of our model. Instead, we perform a random sampling focusing on the regions of interest, the mitochondria. Our pipeline (Supplementary Fig. 1) is described as follows.

- *Anscombe transform.* To detect the areas of interest, we need to enhance the mitochondria signal with respect to the noise. To do this, we first remove the Poisson-Gaussian noise in STED images. This is achieved by applying an Anscombe transform, which enables to stabilize noise variance and to approximately convert Poisson-Gaussian noise into white Gaussian noise with a constant variance. The Anscombe transform of an image I is given by

$$I_{Ansc}(i,j) = 2\sqrt{\frac{3}{8} + I(i,j)}, \forall (i,j) \in \Omega\,. \tag{3}$$

- *Z-score.* Then, we compute the Z-score map defined as

$$Z(i,j) = \frac{I_{Ansc}(i,j) \;-\; \hat{\mu}_{\in}}{\hat{\sigma}_{\in}}, \forall (i,j) \in \Omega\,, \tag{4}$$

where $\hat{\mu}_{\epsilon}$ and $\hat{\sigma}_{\epsilon}$ are the estimated mean and standard deviation of the Gaussian noise ϵ, respectively. Since most of the pixels in I belong to the background, we consider $\hat{\mu}_{\epsilon} =$

$median(\{I(i,j)\}_{(i,j)\epsilon\Omega})$. For $\hat{\sigma}_\epsilon$, we use a robust estimator derived from the Median Absolute Deviation (*MAD*) such that $\hat{\sigma}_\epsilon = 1.4826 \cdot median(\{\,|r(i,j)|\}_{(i,j)\epsilon\Omega})$, where $r(i,j) = \frac{2I_{Ansc}(i,j) - I_{Ansc}(i+1,j) - I_{Ansc}(i,j+1)}{\sqrt{6}}$, $\forall(i,j)\epsilon\Omega$, are the pseudo-residuals. In fact, under the hypothesis of having a white Gaussian noise and that the noise-free image is piecewise smooth in a local neighborhood, we have that $\hat{\sigma}_\epsilon^2 = E[r^2(i,j)]$.

- *Thresholding.* The higher the Z-score in Eq. (4), the higher the pixel value is above the mean of the measured noise and therefore the pixel (i,j) is considered as a pixel of interest. We apply a threshold *c*, in a way that any pixel $(i,j)\epsilon\,\Omega$ such that $Z(i,j) > c$ is considered as a mitochondria pixel. We denote this set as Ω_{mito}. The threshold is automatically adapted for each training image. Starting from a fixed high value of 30, while Ω_{mito} does not contain a minimum of 10% mitochondrial information (i.e., $\#\Omega_{mito}$ < 10% $\#\Omega$, where $\#\Omega$ and $\#\Omega_{mito}$ denote the number of pixels in the sets Ω and Ω_{mito}, respectively), we subtract 5 from the threshold value. This creates a binary mask on which we apply a median to remove the surrounding noise. This automatic procedure avoids cumbersome manual annotations. Note that this mask can also be used to compute the parameter $\gamma_{i,j}$ in our loss (see Eq. (1)).

- *ROIs selection.* From Ω_{mito}, we randomly choose N_I different pixels to be the center of ROIs of size 128×128 pixels. Thus, the more pixels of mitochondria a ROI contains, the more likely it is to be chosen. The following conditions have to be respected: i) the ROI centers should not belong to the borders of the image; ii) to avoid redundancy, a minimum distance of 60 pixels is established between each pairwise ROI center. The resulting ROIs are finally used to create the patches from the normalized training data (see above “Data normalization”).

Network evaluation. In addition to a quantitative comparison to the state-of-the-art methods and experiments to show the reliability of our method (Results), an ablation study was also performed (Supplementary Fig. 4; Supplementary Note 2.3) to highlight the individual contribution of key components of our method. More details about the evaluation metrics and the implementation details of DeepCristae are also given in Supplementary Note 2.1 and Supplementary Note 2.2.1, respectively.

Other Methods and Materials

Cell culture and biological materials, fluorescence labeling, all used microscopy techniques, image acquisition protocols and quantitative measurements are detailed in the Supplementary Note 1. PKMITO dyes are commercially available at Spirochrome (Stein-am-Rhein, Switzerland) and Genvivo Biotech (Nanjing, China)

Data availability

Data can be accessed in the following private link: https://figshare.com/s/39819e2b1a84a7afdb7d

Code availability

DeepCristae source code used in this publication is open-source and published under the BSD 3-Clause "Original" or "Old" License. Source code will be available shortly through GitHub.

References

1. Bock, F. J. & Tait, S. W. G. Mitochondria as multifaceted regulators of cell death. *Nat Rev Mol Cell Biol* **21**, 85–100 (2020).
2. Stimpfel, M., Jancar, N. & Virant-Klun, I. New Challenge: Mitochondrial Epigenetics? *Stem Cell Rev Rep* **14**, 13–26 (2018).
3. Chakrabarty, S. *et al.* Mitochondria in health and disease. *Mitochondrion* **43**, 25–29 (2018).
4. Stephan, T., Roesch, A., Riedel, D. & Jakobs, S. Live-cell STED nanoscopy of mitochondrial cristae. *Sci Rep* **9**, 12419 (2019).
5. Jakobs, S., Stephan, T., Ilgen, P. & Brüser, C. Light Microscopy of Mitochondria at the Nanoscale. *Annu Rev Biophys* **49**, 289–308 (2020).
6. Ishigaki, M. *et al.* STED super-resolution imaging of mitochondria labeled with TMRM in living cells. *Mitochondrion* **28**, 79–87 (2016).
7. Yang, Z. *et al.* Cyclooctatetraene-conjugated cyanine mitochondrial probes minimize phototoxicity in fluorescence and nanoscopic imaging. *Chem Sci* **11**, 8506–8516 (2020).
8. Vicidomini, G., Bianchini, P. & Diaspro, A. STED super-resolved microscopy. *Nat Methods* **15**, 173–182 (2018).
9. Yang, X. *et al.* Mitochondrial dynamics quantitatively revealed by STED nanoscopy with an enhanced squaraine variant probe. *Nat Commun* **11**, 3699 (2020).
10. Kilian, N. *et al.* Assessing photodamage in live-cell STED microscopy. *Nat Methods* **15**, 755–756 (2018).
11. Liu, T. *et al.* Multi-color live-cell STED nanoscopy of mitochondria with a gentle inner membrane stain. *Proceedings of the National Academy of Sciences* **119**, e2215799119 (2022).
12. Wang, C. *et al.* A photostable fluorescent marker for the superresolution live imaging of the dynamic structure of the mitochondrial cristae. *Proceedings of the National Academy of Sciences* **116**, 15817–15822 (2019).
13. Zhang, J. *et al.* Low-Power Two-Color Stimulated Emission Depletion Microscopy for Live Cell Imaging. *Biosensors (Basel)* **11**, (2021).
14. Guo, M. *et al.* Rapid image deconvolution and multiview fusion for optical microscopy. *Nat Biotechnol* **38**, 1337–1346 (2020).
15. Koho, S. *et al.* Fourier ring correlation simplifies image restoration in fluorescence microscopy. *Nat Commun* **10**, 3103 (2019).
16. Huang, X. *et al.* Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy. *Nat Biotechnol* **36**, 451–459 (2018).
17. Zhao, W. *et al.* Sparse deconvolution improves the resolution of live-cell super-resolution fluorescence microscopy. *Nat Biotechnol* **40**, 606–617 (2022).
18. Arigovindan, M. *et al.* High-resolution restoration of 3D structures from widefield images with extreme low signal-to-noise-ratio. *Proceedings of the National Academy of Sciences* **110**, 17344–17349 (2013).
19. Weigert, M. *et al.* Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat Methods* **15**, 1090–1097 (2018).
20. Jin, L. *et al.* Deep learning enables structured illumination microscopy with low light levels and enhanced speed. *Nat Commun* **11**, 1934 (2020).
21. Qiao, C. *et al.* Rationalized deep learning super-resolution microscopy for sustained live imaging of rapid subcellular processes. *Nat Biotechnol* **41**, 367–377 (2023).

22. Li, Y. *et al.* Incorporating the image formation process into deep learning improves network performance. *Nat Methods* **19**, 1427–1437 (2022).
23. Boulanger, J. *et al.* Patch-Based Nonlocal Functional for Denoising Fluorescence Microscopy Image Sequences. *IEEE Trans Med Imaging* **29**, 442–454 (2010).
24. Kempen, G. M. P. van, Voort, H. T. M. van der, Bauman, J. G. J. & Strasters, K. C. Comparing maximum likelihood estimation and constrained Tikhonov-Miller restoration. *IEEE Engineering in Medicine and Biology Magazine* **15**, 76–83 (1996).
25. Van Der Voort, H. T. M. & Strasters, K. C. Restoration of confocal images for quantitative image analysis. *J Microsc* **178**, 165–181 (1995).
26. Van Kempen, G. M. P., Van Vliet, L. J., Verveer, P. J. & Van Der Voort, H. T. M. A quantitative comparison of image restoration methods for confocal microscopy. *J Microsc* **185**, 354–365 (1997).
27. Buades, A., Coll, B. & Morel, J. M. A non-local algorithm for image denoising. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* vol. 2 60–65 vol. 2 (2005).
28. Dabov, K., Foi, A., Katkovnik, V. & Egiazarian, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing* **16**, 2080–2095 (2007).
29. Richardson, W. H. Bayesian-based iterative method of image restoration. *JoSA* **62**, 55–59 (1972).
30. Lucy, L. B. An iterative technique for the rectification of observed distributions. *Astron J* **79**, 745 (1974).
31. Gonzalez, R. C. & Woods, R. E. *Digital Image Processing*. (Pearson Education, Inc., 2003).
32. Prigent, S. *et al.* SPITFIR(e): a supermaneuverable algorithm for fast denoising and deconvolution of 3D fluorescence microscopy images and videos. *Sci Rep* **13**, 1489 (2023).
33. Krull, A., Buchholz, T.-O. & Jug, F. Noise2Void - Learning denoising from single noisy images. in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 2124–2132 (2019).
34. Batson, J. & Royer, L. Noise2self: Blind denoising by self-supervision. in *International Conference on Machine Learning* 524–533 (2019).
35. Ledig, C. *et al.* Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 105–114 (2017).
36. Zhang, K., Zuo, W., Chen, Y., Meng, D. & Zhang, L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing* **26**, 3142–3155 (2017).
37. Zhang, Y. *et al.* Image Super-Resolution Using Very Deep Residual Channel Attention Networks. in *Computer Vision – ECCV 2018* (eds. Ferrari, V., Hebert, M., Sminchisescu, C. & Weiss, Y.) 294–310 (Springer International Publishing, 2018).
38. Chen, J. *et al.* Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nat Methods* **18**, 678–687 (2021).
39. Chen, X. *et al.* Self-supervised denoising for structured illumination microscopy enables long-term super-resolution live-cell imaging. *bioRxiv* 2023.04.05.535684 (2023).

40. Wang, J., Fan, J., Zhou, B., Huang, X. & Chen, L. Hybrid reconstruction of the physical model with the deep learning that improves structured illumination microscopy. *Advanced Photonics Nexus* **2**, 16012 (2023).
41. Prigent, S. *et al.* BioImageIT: Open-source framework for integration of image data management with analysis. *Nat Methods* **19**, 1328–1330 (2022).
42. Opstad, I. S., Wolfson, D. L., Øie, C. I. & Ahluwalia, B. S. Multi-color imaging of sub-mitochondrial structures in living cells using structured illumination microscopy. **7**, 935–947 (2018).
43. Hayashi, S. & Okada, Y. Ultrafast superresolution fluorescence imaging with spinning disk confocal microscope optics. *Mol Biol Cell* **26**, 1743–1751 (2015).
44. Chen, B.-C. *et al.* Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science (1979)* **346**, 1257998 (2014).
45. Lang, A., John Peter, A. T. & Kornmann, B. ER–mitochondria contact sites in yeast: beyond the myths of ERMES. *Curr Opin Cell Biol* **35**, 7–12 (2015).
46. Vance, J. E. Phospholipid synthesis in a membrane fraction associated with mitochondria. *Journal of Biological Chemistry* **265**, 7248–7256 (1990).
47. Lackner, L. L. The Expanding and Unexpected Functions of Mitochondria Contact Sites. *Trends Cell Biol* **29**, 580–590 (2019).
48. Valm, A. M. *et al.* Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017).
49. Amorim, J. A. *et al.* Mitochondrial and metabolic dysfunction in ageing and age-related diseases. *Nat Rev Endocrinol* **18**, 243–258 (2022).
50. Vance, J. E. MAM (mitochondria-associated membranes) in mammalian cells: Lipids and beyond. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1841**, 595–609 (2014).
51. van Vliet, A. R. & Agostinis, P. Mitochondria-Associated Membranes and ER Stress. in *Coordinating Organismal Physiology Through the Unfolded Protein Response* (eds. Wiseman, R. L. & Haynes, C. M.) 73–102 (Springer International Publishing, 2018).
52. Rowland, A. A., Chitwood, P. J., Phillips, M. J. & Voeltz, G. K. ER Contact Sites Define the Position and Timing of Endosome Fission. *Cell* **159**, 1027–1041 (2014).
53. Kleele, T. *et al.* Distinct fission signatures predict mitochondrial degradation or biogenesis. *Nature* **593**, 435–439 (2021).
54. Stoldt, S. *et al.* Spatial orchestration of mitochondrial translation and OXPHOS complex assembly. *Nat Cell Biol* **20**, 528–534 (2018).
55. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* 234–241 (2015).

Acknowledgements

This work was supported by the France-BioImaging Infrastructure (French National Research Agency, ANR-10-INBS-04-07, “Investments for the future”) and the Labex Cell(n)Scale (ANR-11-LABX-0038) as part of the Idex PSL (ANR-10-IDEX-0001-02). We acknowledge the Cell and Tissue Imaging (PICT IBiSA, Institut Curie) and the IMACHEM (Collège de France) platforms, also members of the national infrastructure France-BioImaging (ANR-10-INBS-04-01) for access to and maintaining the spinning-disk, Airyscan and STED microscopes. We also wish to thank M. Maurin from Inserm U932 for his help in a preliminary study on STED image acquisition.

Author Contributions

A.B., L.L., C-A.V-C and J.S. conceived the project. S.P. and A.B. designed the framework of DeepCristae, and conducted benchmarks and every experiment relative to the reliability of the network. S.P. implemented the code of DeepCristae and the Jupyter notebooks. L.L., J.S., and C-A.V-C. designed the biological experiments. L.L., and J.S. prepared samples. L.L. with the aid of J.D. performed the acquisitions. L.L. and C-A.V-C analyzed and organized the biological data. L.L., S.P., J.S. and A.B. prepared figures and videos. T.L. and Z.C. provided the mitochondrial dye. A.B., S.P., J.S., L.L. and C-A.V-C. wrote the manuscript. All authors critically discussed the results and commented on the manuscript. A.B., J.S. and C.K. supervised the research.

Competing interests

Z.C. is an inventor of the patent on the mitochondria dye described in this work. All other authors have no competing interests to declare.

Figures

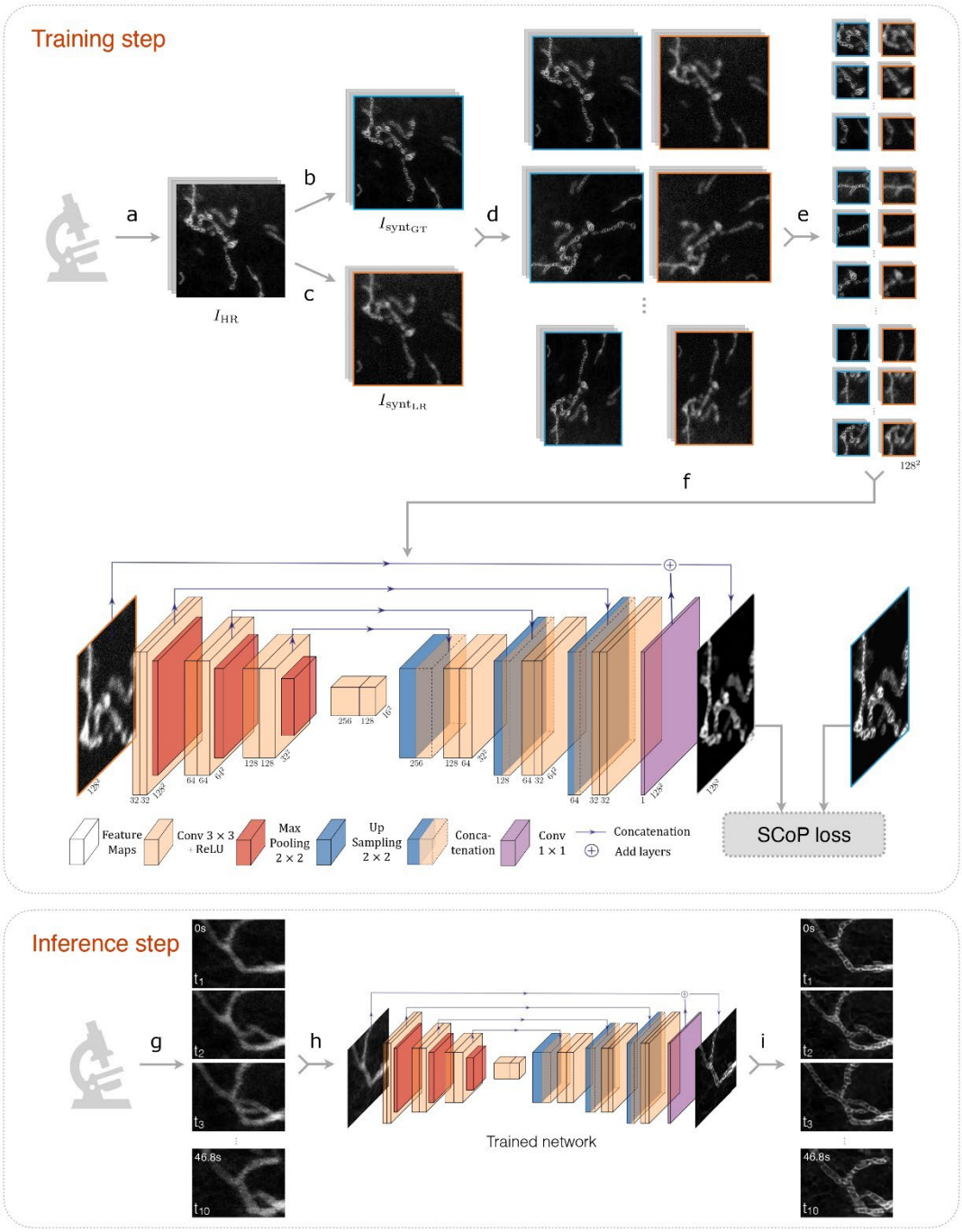
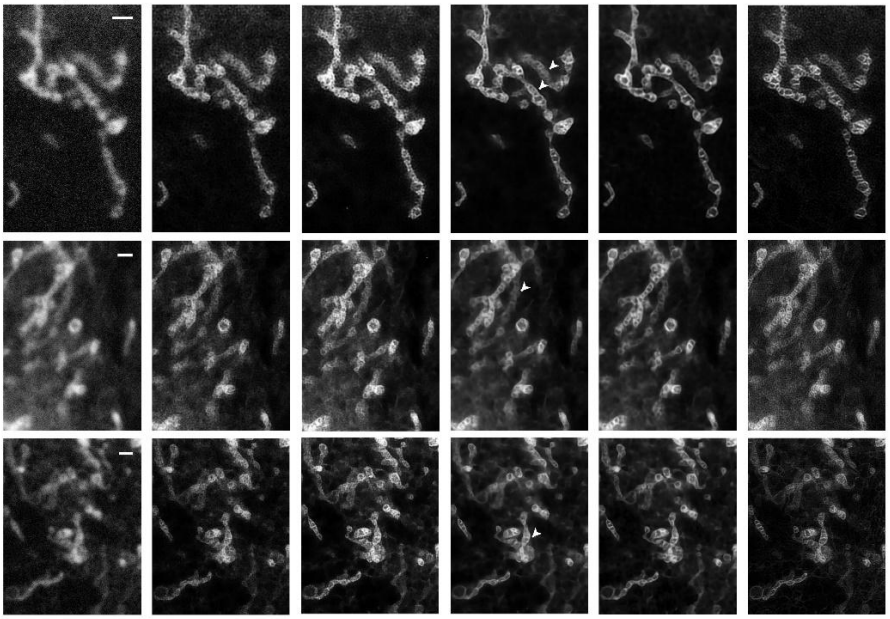


Figure 1. Overview of DeepCristae. *Training step:* **(a)** Acquisition of 33 high-resolution (HR) 2D STED images of live RPE1 cells stained with PKMITO-Orange, for mitochondria. From these HR images I_{HR} a dataset D_{synt} is obtained**: (b)** deconvolution of the HR images using a Richardson-Lucy algorithm to enhance the mitochondria cristae. The resulting images $I_{synt_{GT}}$ are used as ground truths to train the network; **(c)** resolution degradation of the images I_{HR} by applying Gaussian filtering and by corrupting images with Poisson-Gaussian noise. The resulting images are denoted $I_{synt_{LR}}$. The obtained dataset D_{synt} is divided into 24 training

images and 9 test images. To increase the size of the training set, the images $I_{synt_{LR}}$ and ground truth $I_{synt_{GT}}$ are then augmented **(d)** and sampled in patches of size 128x128 pixels **(e)**. We finally obtained 1824 pairs of ground truths (blue) and low-resolution input images (orange) to train our network **(f)**. The training is performed by minimizing our SCoP loss, especially dedicated to restoring mitochondria pixels. *Inference step:* **(g)** Long-term and fast acquisition with low illumination of live mitochondria. Note that if the training was performed on degraded STED images, the inference can be made on other microscopy modalities (e.g., STED, Live-SR and LLSM). **(h, i)** Frame-by-frame restoration of the acquired sequence by our previously trained DeepCristae network, allowing observation of the mitochondrial cristae dynamics at high resolution.

a



b

	PSNR	$PSNR_{mito}$	$PSNR_{cristae}$	SSIM	$SSIM_{mito}$	$SSIM_{cristae}$
DeepCristae	**30.43 ± 2.74**	**24.38 ± 2.80**	**22.04 ± 2.57**	0.72 ± 0.08	**0.58 ± 0.08**	**0.52 ± 0.07**
RL	24.32 ± 3.43	18.97 ± 2.99	19.92 ± 3.39	0.45 ± 0.17	0.43 ± 0.11	0.37 ± 0.08
Wiener	23.89 ± 3.32	19.08 ± 2.80	19.95 ± 3.16	0.38 ± 0.16	0.38 ± 0.11	0.34 ± 0.08
SPITFIR(e)	25.57 ± 1.46	21.29 ± 2.04	20.11 ± 2.16	0.34 ± 0.06	0.48 ± 0.07	0.40 ± 0.06
ND-SAFIR	24.74 ± 3.28	17.88 ± 3.12	19.33 ± 3.83	0.70 ± 0.11	0.49 ± 0.09	0.39 ± 0.06
Median 2D	23.59 ± 3.66	17.31 ± 3.28	18.88 ± 3.95	0.63 ± 0.14	0.47 ± 0.10	0.38 ± 0.06
Noise2Void	23.95 ± 3.63	17.58 ± 3.24	19.07± 3.92	0.64 ± 0.13	0.48 ± 0.10	0.39 ± 0.07
CARE	29.97 ± 2.87	23.50 ± 2.91	21.87 ± 2.73	**0.74 ± 0.10**	0.56 ± 0.10	0.48 ± 0.09
RCAN	28.17 ± 3.53	22.20 ± 3.27	20.56 ± 2.40	0.66 ± 0.14	0.47 ± 0.10	0.41 ± 0.06
SRResNet	23.18 ± 3.54	15.92 ± 3.27	14.78 ± 3.57	0.66 ± 0.12	0.41 ± 0.11	0.38 ± 0.07
Inputs	23.71 ± 3.28	18.72 ± 2.70	19.94 ± 3.16	0.37 ± 0.17	0.37 ± 0.11	0.32 ± 0.07

	NRMSE	$NRMSE_{mito}$	$NRMSE_{cristae}$
DeepCristae	0.029 ± 0.007	**0.057 ± 0.015**	**0.080 ± 0.026**
RL	0.036 ± 0.010	0.065 ± 0.018	0.103 ± 0.037
Wiener	0.040 ± 0.011	0.071 ± 0.021	0.112 ± 0.042
SPITFIR(e)	0.034 ± 0.004	0.064 ± 0.011	0.095 ± 0.025
ND-SAFIR	0.030 ± 0.007	0.061 ± 0.014	0.094 ± 0.028
Median 2D	0.031 ± 0.008	0.062 ± 0.014	0.097 ± 0.030
Noise2Void	0.030 ± 0.007	0.061 ± 0.014	0.094 ± 0.028
CARE	**0.028 ± 0.008**	0.057 ± 0.017	0.083 ± 0.029
RCAN	0.033 ± 0.008	0.066 ± 0.016	0.091 ± 0.029
SRResNet	0.032 ± 0.007	0.066 ± 0.014	0.093 ± 0.028
Inputs	0.040 ± 0.011	0.071 ± 0.021	0.114 ± 0.042

Figure 2. DeepCristae outperforms state-of-the-art methods for restoring mitochondria cristae in low-resolution 2D STED images. (a) The image grid displays restoration results of 3 test images from the dataset D_{synt} obtained with DeepCristae and three competitive deep learning methods: RCAN[37], SRResNet[35] and CARE[19]. GT: ground truth. Pixel size: 25 nm. Scale bar: 1 µm. White arrowheads indicate mitochondria with low contrast or in a noisy background restored by CARE; to be compared with DeepCristae column **(b)** Quantitative comparison of DeepCristae with conventional (Richardson-Lucy (RL)[29,30], Wiener[31], SPITFIR(e)[32], ND-SAFIR[23], 2D median filter (Median 2D) and deep learning (Noise2Void[33], CARE[19], RCAN[37] and SRResNet[35]) image restoration methods. Metrics were computed on the 9 test images of D_{synt}. Note that all deep learning methods were trained using the same patches extracted from the training images of D_{synt}. Parameters used for conventional methods are indicated in Supplementary Note 2.2.

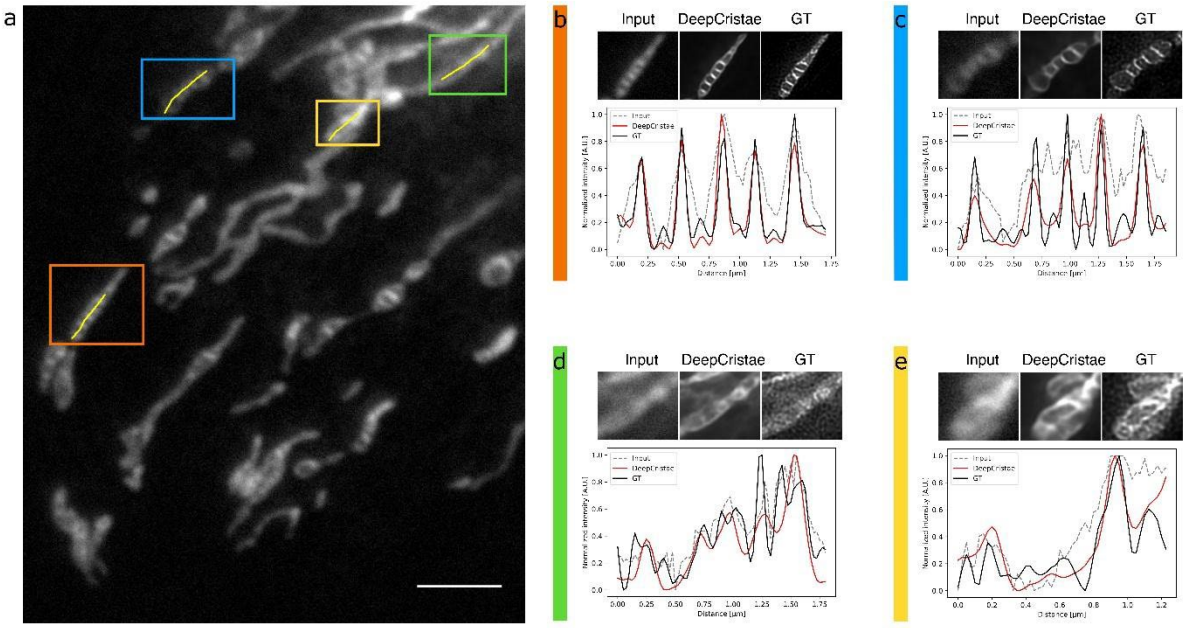
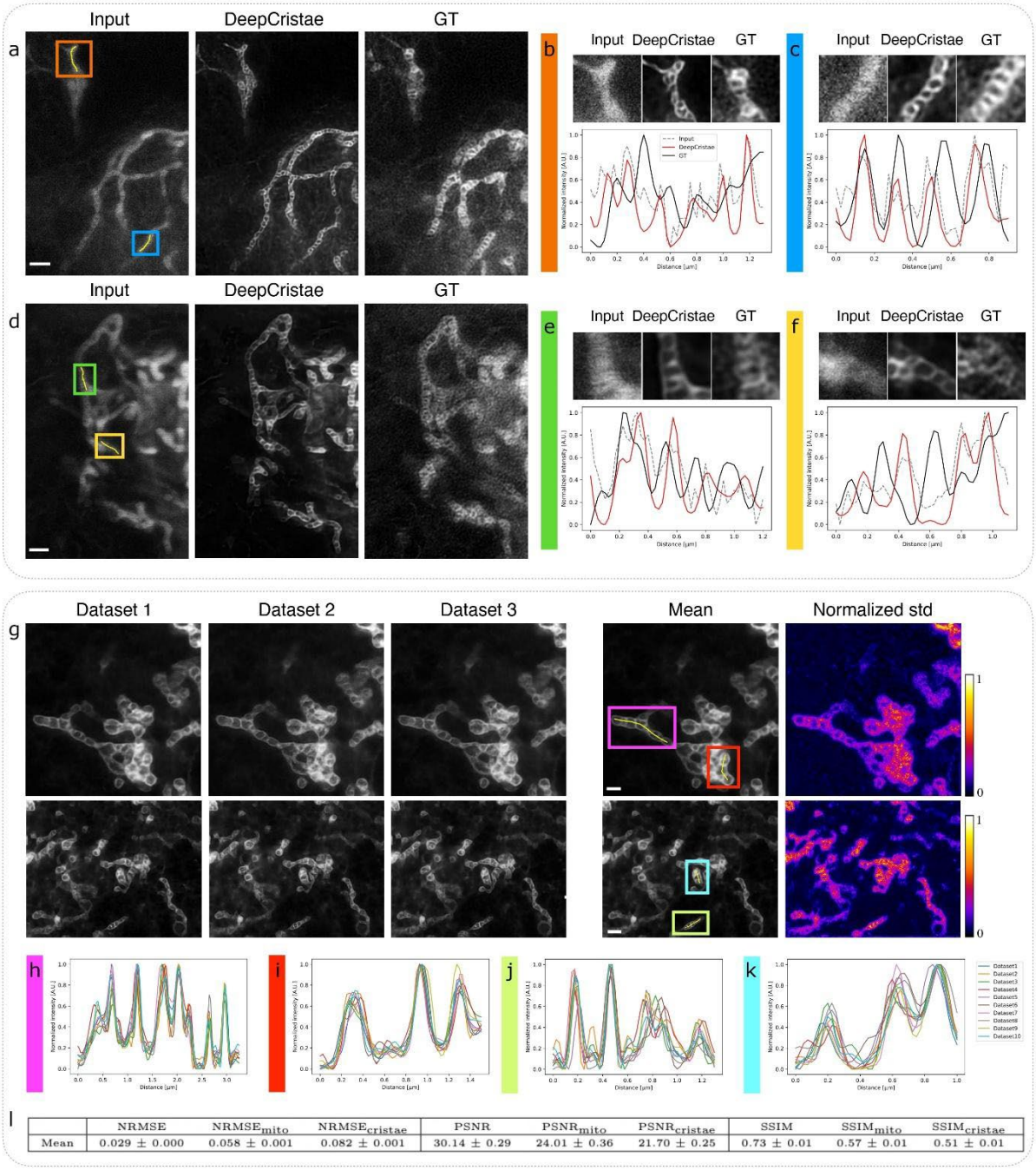


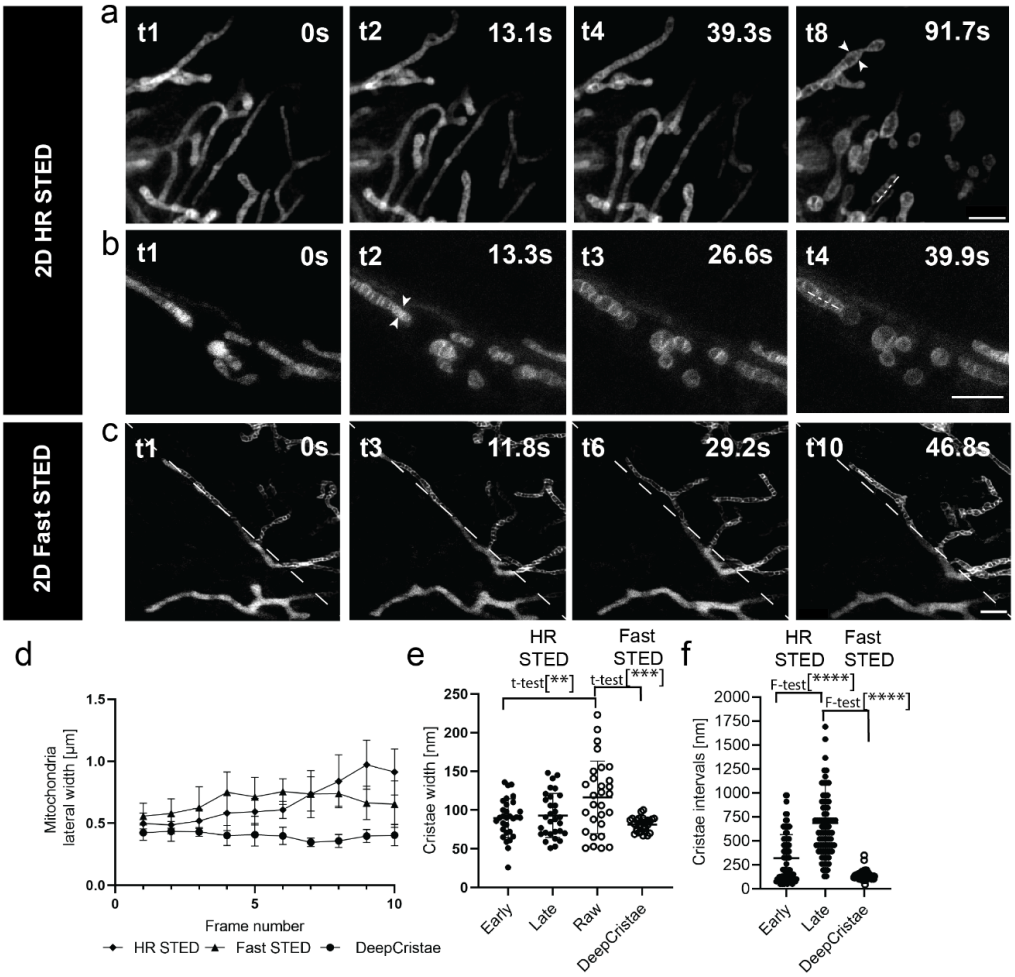
Figure 3. DeepCristae reveals mitochondria cristae from low resolution (LR) 2D STED images. RPE1 cells were labeled with PKMITO-Orange prior to high resolution (HR) 2D STED imaging. **(a)** Degradation of the HR image with a Gaussian filter ($\sigma_{blur} = 3.25$ pixels) and additive Poisson-Gaussian noise ($\sigma_{noise} = 4.0$). The resulting image is part of the test dataset of D_{synt}**.** Pixel size: 25 nm. Scale bar: 3 µm. **(b-e)** Zoom in on the colored insets depicted in **(a)**. Top, from left to right: thumbnails of the zoomed-in area on the LR image, the image restored by DeepCristae and the ground truth (GT) image, respectively. The GT was obtained by deconvolving the HR raw image with the RL algorithm to enhance mitochondria cristae. Bottom: comparison of normalized intensity line profiles along a mitochondrion between the three thumbnails. The yellow line indicated in the corresponding colored inset in **(a)** serves to identify the fluorescence profile. Note that GT shows undesired high-intensity pixels in the background due to noise amplification by the RL algorithm. Therefore, DeepCristae line profiles are sometimes smoother than the one of the ground truth, especially in **(d)** and **(e)**.



	NRMSE	$NRMSE_{mito}$	$NRMSE_{cristae}$	PSNR	$PSNR_{mito}$	$PSNR_{cristae}$	SSIM	$SSIM_{mito}$	$SSIM_{cristae}$
Mean	0.029 ± 0.000	0.058 ± 0.001	0.082 ± 0.001	30.14 ± 0.29	24.01 ± 0.36	21.70 ± 0.25	0.73 ± 0.01	0.57 ± 0.01	0.51 ± 0.01

Figure 4. Reliability of image restoration by DeepCristae. *First experiment* ***(a-f)*****:** assessment of the reliability of DeepCristae on real data by controlling the consistency between DeepCristae restoration and a real "ground truth". RPE1 cells were labeled with PKMITO-Orange prior to low resolution (LR) and to high resolution (HR) 2D STED imaging, acquired successively. **(a, d)** Two LR STED images (Inputs) were given as input to DeepCristae for restoration and qualitatively compared to the corresponding HR STED images. Notes that the HR data were deconvolved using the Richardson-Lucy algorithm to enhance mitochondria cristae and are considered here as "ground truths" (GTs). Pixel size: 25 nm. Scale bar: 1 µm. **(b, c) and (e,f)** Zoom in on the colored insets depicted in **(a)** and **(d)**, respectively. Top, from left to right: thumbnails of the zoomed-in area on the LR image, the image restored by DeepCristae and the ground truth (GT) image, respectively. Bottom:

comparison of normalized intensity line profiles along a mitochondrion between the three thumbnails. The yellow line indicated in the corresponding colored inset in **(a)** and **(d)** serves to identify the fluorescence profile. Note that the line profiles were taken in regions where small mitochondrial displacements were observed. Despite an offset due to mitochondrial motion, a consistency between the cristae restored by DeepCristae and the ones present in the ground truths is observed overall. *Second experiment* **(g-l)**: assessment of the reliability of DeepCristae by studying the consistency between its predictions obtained with different training. To that end, 10 DeepCristae neural networks were trained with different training data, each one generated with our patch generation method applied to the 24 training images of D_{synt}. Note that for this experiment, all networks were initialized with the same weights. **(g)** From left to right: predictions of three DeepCristae networks on two test images of D_{synt}, the average prediction over the 10 trainings and the corresponding pixel-wise normalized standard deviation. Pixel size: 25 nm. Scale bar: 1 μm. **(h-k)** Comparison of normalized intensity line profiles along a mitochondrion in **(g)** between the 10 trainings. The yellow line indicated in the corresponding colored inset in **(g)** serves to identify the fluorescence profile. **(l)** Quantitative comparison of the 10 DeepCristae models. Metrics were computed on the 9 test images of D_{synt}.



g

	Cristae width		Cristae intervals	
	mean [nm]	SD [nm]	mean [nm]	SD [nm]
Early	90	25	319	247
Late	93	28	687	417
Raw	116	47		
DeepCristae	81	9	142	46
	t-test p-val	F-test p-val	t-test p-val	F-test p-val
Early-Late	ns	ns	<0.0001[****]	<0.0001[****]
Early-Raw	0.0082[**]	0.0012[**]		
Early-DeepCristae	ns	<0.0001[****]	<0.0001[****]	<0.0001[****]
Late-Raw	0.0242[*]	0.0089[**]		
Late-DeepCristae	0.0378[*]	<0.0001[****]	<0.0001[****]	<0.0001[****]
Raw-DeepCristae	0.0003[***]	<0.0001[****]		

Figure 5. DeepCristae reveals mitochondria cristae from low resolution 2D live STED. (a-c) 2D live STED imaging of RPE1 cells labeled with PKMITO-Orange. **(a, b)** Time series of 10 images (Δt~13s) using 2D high-resolution (HR) STED (Supplementary Note 1: Materials and Methods). Pixel size is 50 x 50 nm in (**a)** and 25 x 25 nm in (**b**). Early and late time points are shown in **a**, the 4 first time points are shown in **b.** Phototoxic damage as illustrated by the swelling of the mitochondria. **(c)** Time series of 2D Fast STED, improving the time delay between time points (Δt~5.9 s) and reducing mitochondrial damage. Each thumbnail is diagonally divided for raw 2D Fast (LR) STED images (bottom-left part) and after DeepCristae treatment (top-right). To match the training mitochondria settings (see Results "Reliability and

limits of image restoration by DeepCristae”), a rescaling of 1.87 is applied to the LR image before DeepCristae inference. Scale Bars in **a-c** are: 2µm. **(d)** Lateral widths of 7 mitochondria measured at each time point (series of 10 time points) in HR STED and in Fast (LR) STED, before and after DeepCristae (line profiles as depicted as arrow heads in **a** t8 and **b** t2). Line profiles were fitted to a Gaussian model and FWHM was measured, as indicated in Supplementary Note 1: Materials and Methods. Data are expressed as mean ± SD. **(e)** Cristae widths measured as in **d** for 30 cristae on 20 mitochondria (from 4 distinct image series) in early and late time points of HR STED and 31 cristae in Fast (LR) STED, before (Raw; two measures above the ordinate scale) and after DeepCristae (line profiles as illustrated in **a** t8, **b** t4 and **c** t10). Data are expressed as mean ± SD. Student t-test: ** (p-val=0.0082), ***(p-val=0.0003). **(f)** Distances between two cristae (Cristae intervals) measured in a peak-to-peak intensity from plot profiles (RAW Fast STED was not possible). Early: N=56 from 2 series; Late: N=80 in 2 series; DeepCristae: N=60 in 3 series, distributed in all time points. Data are expressed as mean ± SD. F-test: ****(p-val<0.0001). **(g)** Table with complete statistics including significance using Student- and Fisher-tests; ns= non-significant.

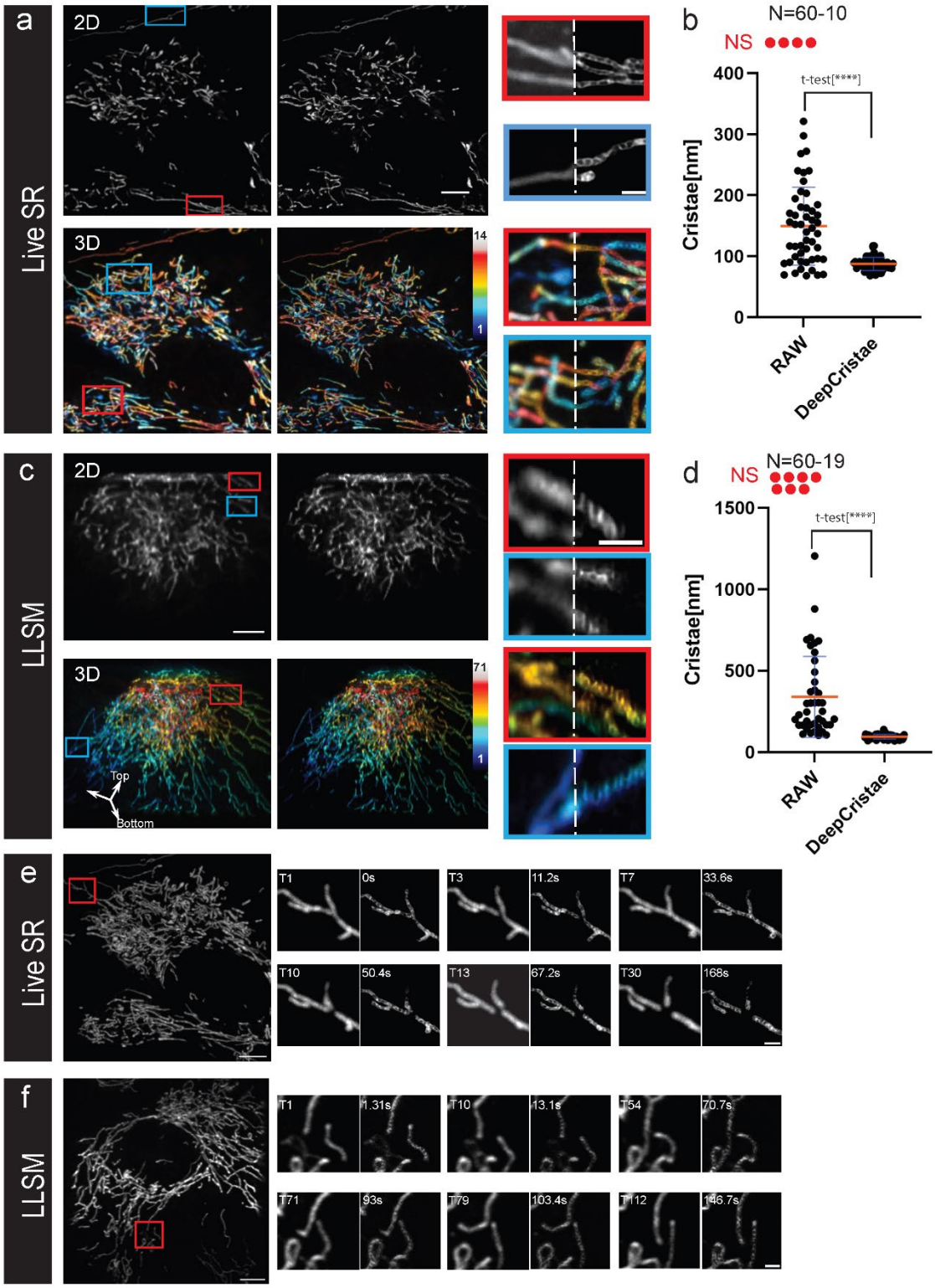


Figure 6. DeepCristae restoration enhances cristae width resolution in 3D and 3D+live imaging. (a) 2D plane (top) and 3D MIP (Maximum Intensity Projection of 14 planes) (bottom) of cell labeled with PKMITO-Orange, acquired with a SD microscope with a Live-SR module, before (left) and after DeepCristae (right). Thumbnails are zoomed areas corresponding to the insets (red and blue) and are composites of raw and DeepCristae images. Color scale bar indicates the bottom to top position of mitochondria (bottom right). **(b)** Cristae widths were

measured as in **Fig. 5e**; each individual measurement in DeepCristae restored images is compared to its equivalent in raw images except for 10 cristae, not measurable in Raw (N=60 and N=60-10, respectively). Data are expressed as mean ± SD (DeepCristae=87 ± 1 nm; Raw =149 ± 64 nm; Student t-test, [****] p<0.0001). **(c)** One section plane (top) and 3D reconstructed MIP of 71 planes (bottom) of a cell acquired with a Lattice Light Sheet microscope (LLSM) in a dithered mode, before (left) and after DeepCristae (right). Thumbnails are zoomed areas corresponding to the insets (red and blue) and are composite of raw and Deepcristae restored images. Color scale bar indicates the bottom to top position of mitochondria (bottom right). **(d)** Cristae widths were measured as in **Fig. 5e** in DeepCristae restored images and compared to its Raw equivalents, when possible (N=60 and N=60-19, respectively). Data are expressed as mean ± SD (DeepCristae=94 ± 15 nm; Raw=339 ± 248 nm; Student t-test, [****] p<0.0001). **(e, f)** 3D+time imaging using Live-SR **(e)** or LLSM **(f)**. MIP of single time points are shown (left images). Insets indicated in red are zoomed in the thumbnails (right image series) to illustrate fusion or fission dynamics of mitochondria. The selected zoomed areas are shown at different time points before (left panel) and after (right panel) DeepCristae restoration. Time frames between stacks are 5.6s and 1.31s in double channel acquisition for Live-SR and LLSM, respectively. Scale bars = 5µm in full field image and =1µm in zoomed thumbnails. Before DeepCristae restoration, rescaling of 2.6 and of 4.16 were first applied to each raw Live-SR and LLSM data, respectively (see Results “Reliability and limits of image restoration by DeepCristae”).

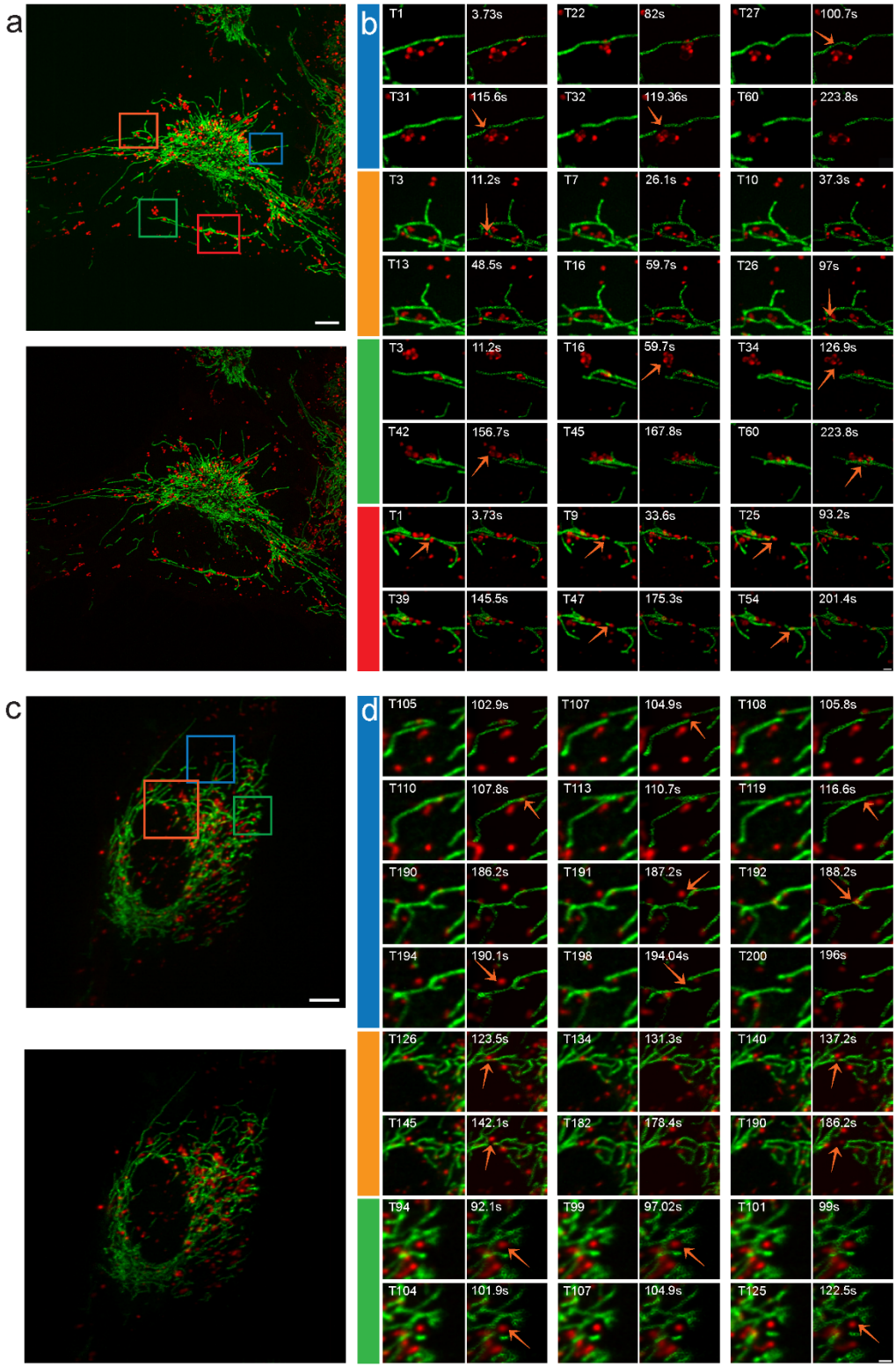


Figure 7. DeepCristae reveals 3D+time cristae morphology during endo/lysosome mitochondria interactions. RPE1 cells incubated 4 hours with Cell Mask PM Deep Red (red) were labeled with PKMITO-Orange (green) for the last 15 minutes. **(a)** A Maximum Intensity Projection (20 planes; Stack time=1.86s/channel, time point t1 over 60) image acquired with Live-SR microscopy is shown before (top) and after (bottom) DeepCristae restoration of the

mitochondria (green channel) as well as after denoising (ND-SAFIR) and RL deconvolution of the endo/lysosomes (red channel). Colored Insets indicate intracellular locations with dynamic events of interest. **(b)** Thumbnails show selected time points of zoomed area as indicated by insets in **a.** They are presented as paired images before (left panels) and after (right panels) DeepCristae restoration. Both time points (left panels) and time frames in seconds (right panels) are indicated for comparison with **d**. **(c)** 3D reconstruction (56 planes; stack time=0.49s/channel, time point t1 over 200) of data acquired with LLSM after deskewing and RL deconvolution is shown before (top) and after (bottom) DeepCristae restoration (green channel). Colored Insets indicate intracellular locations with a focus on dynamic events of interest. **(d)** shows thumbnails of zoomed areas indicated by insets in **c**. Note that for the "blue" area, an early time (T105-T119) and a late (T190-T200) time series are represented, showing mitochondria fusion and fission events, respectively. Orange arrows focus on local interactions between endo/lysosomes and mitochondria**.** Scale bars in **(a, c)** and **(b, d)** are equal to 5µm and 1µm, respectively. Note that in these experiments, before DeepCristae restoration, rescalings of 2.6 and of 4.16 were first applied to each raw Live-SR and LLSM data, respectively, to match the training mitochondria settings (see Results "Reliability and limits of image restoration by DeepCristae".