nature methods

Article

https://doi.org/10.1038/s41592-023-01815-0

ERnet: a tool for the semantic segmentation and quantitative analysis of endoplasmic reticulum topology

Received: 19 May 2022

Accepted: 10 February 2023

Published online: 30 March 2023

Check for updates

Meng Lu[1,2], Charles N. Christensen[1,3], Jana M. Weber[1,5], Tasuku Konno[4], Nino F. Läubli[1], Katharina M. Scherer[1], Edward Avezov[4], Pietro Lio[3], Alexei A. Lapkin[1], Gabriele S. Kaminski Schierle[1,2] & Clemens F. Kaminski[1,2]

The ability to quantify structural changes of the endoplasmic reticulum (ER) is crucial for understanding the structure and function of this organelle. However, the rapid movement and complex topology of ER networks make this challenging. Here, we construct a state-of-the-art semantic segmentation method that we call ERnet for the automatic classification of sheet and tubular ER domains inside individual cells. Data are skeletonized and represented by connectivity graphs, enabling precise and efficient quantification of network connectivity. ERnet generates metrics on topology and integrity of ER structures and quantifies structural change in response to genetic or metabolic manipulation. We validate ERnet using data obtained by various ER-imaging methods from different cell types as well as ground truth images of synthetic ER structures. ERnet can be deployed in an automatic high-throughput and unbiased fashion and identifies subtle changes in ER phenotypes that may inform on disease progression and response to therapy.

The ER is the largest membranous structure in eukaryotic cells and acts as a platform for protein synthesis and quality control and for various organelle interactions[1]. The ER consists of distinct domains including sheets and tubules and features growth tips and tubular connections, so-called three-way junctions. Perturbations to the ER structure and dynamics caused by genetic defects or metabolic stress have been associated with a variety of diseases[2], such as hereditary spastic paraplegias (HSPs) and Niemann–Pick disease type C (NPC). Hence, to understand the role of the ER in diseases, it is important and necessary to characterize ER morphology comprehensively, which may provide powerful phenotypes to screen drugs against ER-associated disorders. However, given the extent of the ER network and its complexity, precise and quantitative measurement of ER topology and movement has remained challenging.

The ER network in a single cell consists of thousands of interconnected tubules that undergo constant rearrangements via processes including continuous tubular elongation, contraction and fusion. Furthermore, there are rapid transitions between sheet and tubular domains with distinct putative functions[3]. Recently, capabilities have emerged to reveal such dynamic changes in ER topology in live cells at subwavelength resolution[4]. Structured illumination microscopy (SIM), for example, can be used to resolve details of ER topology and its rapid remodeling process[5–7]. However, the data have only been interpreted qualitatively, without attempts to quantify ER topology or its structural

[1]Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, UK. [2]Cambridge Infinitus Research Centre, University of Cambridge, Cambridge, UK. [3]Artificial Intelligence Group, Department of Computer Science and Technology, University of Cambridge, Cambridge, UK. [4]UK Dementia Research Institute at the University of Cambridge and Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK. [5]Present address: Delft Bioinformatics Lab, Intelligent Systems Department, Delft University of Technology, Delft, the Netherlands.
e-mail: cfk23@cam.ac.uk

changes precisely. Compared to other organelles, such as mitochondria and lysosomes, which are structurally simpler organelles that are often well separated from one another, the ER consists of highly convoluted and structurally connected domains. The task is further complicated by the fact that the signal-to-noise ratio (SNR) of images obtained during live cell microscopy is often poor, while clear differentiation of the organelle from its background is required to ensure successful segmentation into tubular and sheet domains. For moving structures and time-lapse imaging, this becomes a formidable task.

A number of machine learning-based methods have been developed for the segmentation of cells[8], mitochondria[9,10] and nuclei[11], which provide robust and precise classification of cell structures. However, to date, thresholding remains the standard method of use for ER segmentation[12–14]. Thresholding lacks both sensitivity and specificity, making quantitative conclusions hard to draw, especially in situations in which image quality is compromised by noise. Alternative methods are based on labor-intensive manual labeling of image data to generate specialized datasets for training of machine learning algorithms. These approaches do not generalize well to work with changing experimental setups or varying sample types[15] (Extended Data Fig. 1). An additional challenge for ER segmentation can be seen in temporal consistency. Conventional segmentation is performed on a frame-by-frame basis, and segmented structures in sequential (time-lapse) images lose temporal continuity and thereby cause artifacts[16]. Currently, there is no ER-segmentation method capable of taking dynamic, spatial and temporal topology changes into consideration.

To address these difficulties, we developed ERnet, a deep learning software that automatically segments the ER, classifies its domains into tubules and sheets and quantifies structural and dynamic features in image sequences obtained from live cells. ERnet is trained with image datasets to model the domain knowledge of ER structures, that is, the shapes of tubules and sheets. As a result, it enables feature-specific segmentation with enhanced robustness, specificity and sensitivity regardless of pixel intensity in the images. ERnet works on two-dimensional (2D) data, but the quantitative results accurately describe the three-dimensional (3D) structure of the organelle. After segmentation, ERnet quantifies topological features of the ER and recognizes subtle changes in ER structure and dynamics for various stress conditions, including gene knockout (KO) and knockdown, ATP depletion, calcium depletion, etc. To validate the method, we tested the segmentation accuracy of ERnet on in vitro models subjected to different genetic and metabolic manipulations, including cells mimicking phenotypes of HSP and NPC. Two phenotypes were identified as sensitive readouts of the ER response in these models, namely, the degree of fragmentation of ER networks and the heterogeneity in tubule connections. Both are indicators for the functional state of the ER network and can be used, for example, to quantify the degree of disorganization, shrinkage and collapse of ER structures in models of disease. We show the versatility of ERnet by application to widefield imaging, confocal and super-resolution microscopy data and test its performance in the presence of image noise. Furthermore, the method works in multiple cell lines. Minimal training or no retraining is required between different use scenarios. We provide ERnet as a user-friendly, open-source software package with a graphical user interface (Extended Data Fig. 2 and Supplementary Note) to make it a broadly accessible tool for biologists and to promote ER-related research in basic science and clinical applications.

Results

Design and workflow of ERnet

The general design of ERnet is schematized in Fig. 1a. First, reconstructed sequential images of the ER were segmented in ERnet, followed by classification of ER structures into tubules and sheets. The tubular structure was further skeletonized using a surface axis thinning algorithm[17]. After this, the nodes and edges of the skeletonized ER were identified to plot a topology graph via a graph theory-based module[18]. In essence, the topology graph is a representation of ER tubules and junctions that provides a visual cue on the degrees of ER network connectivity and fragmentation (for an introductory explanation of graph theory concepts, see Extended Data Fig. 6).

Instead of relying on the commonly applied convolutional neural networks (CNNs), our model builds upon a Vision Transformer architecture[19] that outperforms a comparable state-of-the-art CNN with higher classification accuracy and requires four times fewer computational resources. Key to our method is that, rather than paying attention to the spatial position of the nodes, it focuses on the ER's network features, for example, connectivity between nodes. This means that metrics such as number of ER fragments and clustering of nodes into subregions can be extracted to provide quantitative metrics of ER topology and health.

To reduce the computational cost associated with the large data volumes generated by time-sequenced imaging data, ERnet makes use of the backbone architecture of the Swin Transformer reported by Liu et al.[20]. Here, image frames in a temporal sequence are processed as 3D blocks, which permits the model to focus on key features that persist not only over the spatial domain but also over temporal domains (Fig. 1b). These attributes make the method fast to execute and also very responsive to changing ER phenotypes.

Quantitative segmentation and analysis of ER topology

The ER is a highly dynamic structure, and, at any instant, thousands of tubules move and change position, direction and network connections. To quantify these intracellular changes, we first tested the performance of ERnet using SIM images of COS-7 cells. Fig. 2a shows a single frame of the ER (gray) from a set of sequential images captured from a COS-7 cell expressing mEmerald-Sec61b[5]. The performed segmentation successfully identified the whole ER structure, differentiated it from the cytosol background and further classified it into tubular (cyan) and sheet domains (yellow) (Fig. 2a). Next, the tubular ER was skeletonized from the whole structure, and nodes (tubule junctions, shown in red) and edges (tubules, green) were identified as two key topological components to map network connectivity via the Python package graph-tool[18].

SIM provide high spatial–temporal resolution of ER structures and is thus suitable for live cell imaging. A single pixel on the camera frame has a length scale of 42 nm in real space, almost a quarter of the average width of an ER tubule (~160 nm, measured as the average width on SIM images taken). This means that misclassification of a few or even just one image pixel(s) can mean the difference between identification of a tubule as connected or as disrupted. This leads to errors in classification of network features and vice versa to bias when quantifying network connectivity. In disease models, this could lead to erroneous phenotypes. Semantic segmentation of individual pixels from SIM images ensures the structural integrity of networks identified and prevents information loss, an improvement over traditional algorithms used in the past. Fig. 2a,b shows how the method performs. Clear segmentation of ER structure (Fig. 2b) is achieved in regions containing dense ER tubule networks, as can be seen from the enlarged region indicated by the white box in Fig. 2a. This permits the distinction of tubules and their junctions in confined regions, measuring less than 500 nm across (highlighted by yellow dashed lines) with good structural detail. The segmented ER was then skeletonized (Fig. 2a,b, middle) and classified into edges (Fig. 2a,b, right, green tubules) and nodes (Fig. 2a,b, right, red spots). Finally, ERnet quantified the number of edges and nodes (Fig. 2c, top) and the percentage of area covered by tubules and sheets (Fig. 2c, bottom), respectively, across the whole ER. Here, ER tubules were defined as linear branched structures and sheets as flat membrane cisternae as shown in Fig. 2a,d. Morphological features, such as the percentage of tubules or sheets among the whole ER, reflect ER status[3] and provide indications for

https://doi.org/10.1038/s41592-023-01815-0

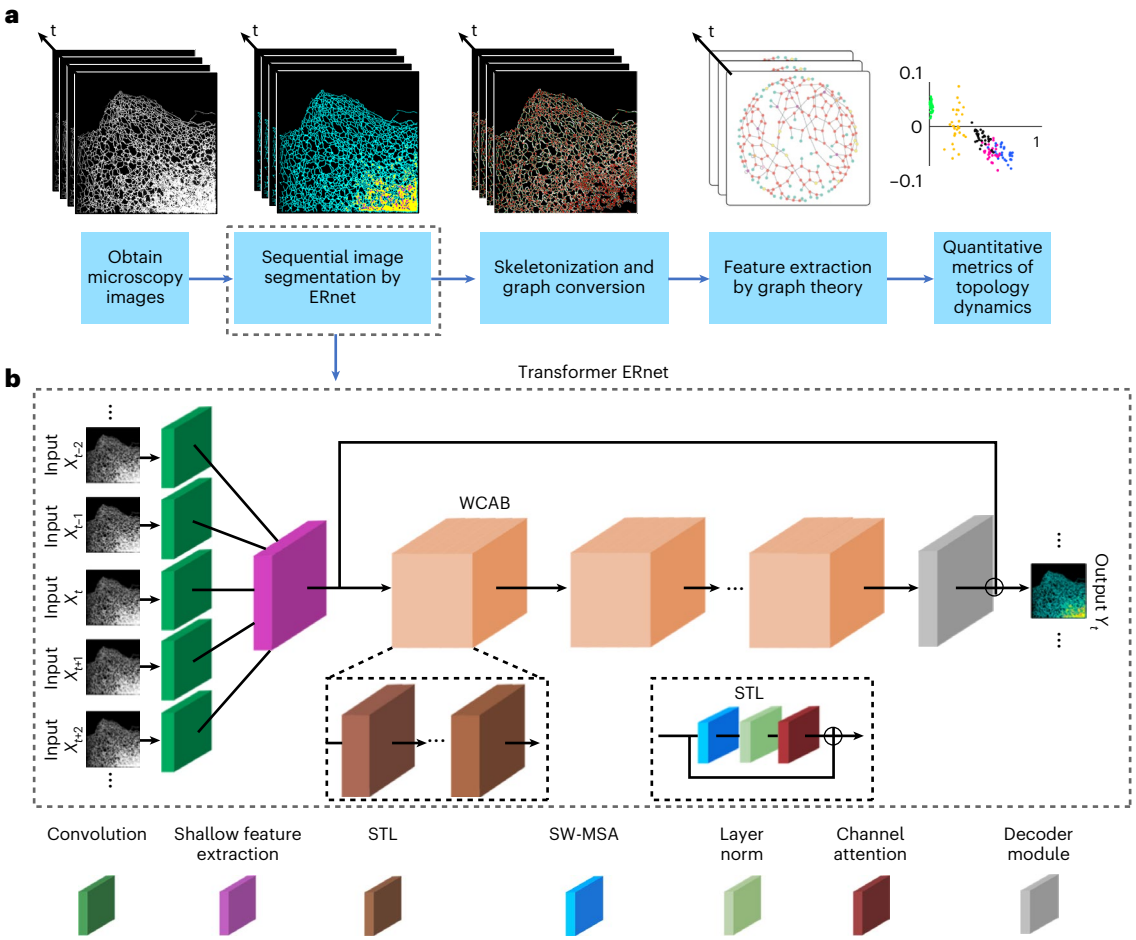


Fig. 1 | Workflow of ER structure segmentation and ERnet construction. **a**, The processing pipeline of ER segmentation and analysis. Time-lapse SIM images were first segmented by ERnet to classify tubules and sheets. The tubular network of the ER after segmentation was further skeletonized, and nodes and edges were identified to plot the connectivity graph. Using graph theory-based methods, we quantified the metrics of ER network features that describe topology and dynamics. **b**, The Transformer-based architecture of ERnet. A moving window loads adjacent frames (X_{t-2} to X_{t+2}) as inputs from time-lapse images into ERnet. A shallow feature-extraction module then projects the input into a feature map, which is followed by a sequence of residual blocks denoted as a window channel attention block (WCAB). Inside each WCAB, there is a sequence of Swin Transformer layers (STLs). For details, see Methods. SW-MSA, shifted window multi-head self-attention.

possible ER defects. ER stress induced by an absence of the GTPase Rab7, which is known to modulate lysosome–ER contact sites, leads to enlargement of ER sheets and reduction of tubular domains in the cell periphery[21]. On the other hand, a depletion of protrudin, an ER-reshaping protein, induces HSP-associated ER dysfunctions by disrupting the sheet-to-tubule balance[22]. Therefore it is expected that topological features of the ER, such as its connectivity, assortativity or clustering coefficients, change for different phenotypes and with disease progression, a topic that is further explored in subsequent sections. It is worth highlighting the fact that, although the ER tubular network underwent stark morphology changes (Supplementary Video 1) and demonstrated fluctuations in the number of nodes and edges (Fig. 2c, top) within individual recordings, its tubule and sheet percentages of the whole ER remained stable (Fig. 2c, bottom), which suggests that the overall connections do not change in the absence of stimuli.

In the canonical model of ER structures, ER tubules radiate from sheets toward the cell periphery[4], and the two structures are thought not to overlap. However, we observed that tubular structures also reside on the ER sheets themselves (Fig. 2d and Supplementary Video 2), which, in what follows, we refer to as 'sheet-based tubules' (SBTs) and are clearly distinguished by ERnet as seen in Fig. 2d and Supplementary Videos 2 and 3. Similar to peripheral tubules, SBTs undergo rapid elongations and contractions, which can lead to either new tubular connections (blue arrows) or separations (gray arrows) (Fig. 2d, bottom). A subsequent 3D reconstruction of SIM image sections further validated that such tubules are directly attached to sheets and are not the result of a projection view artifact (Fig. 2e, Extended Data Fig. 3 and Supplementary Video 4). Analysis of 500 cells showed that this phenomenon is a common feature of the ER network (Fig. 2f).

In silico validation of ERnet

To examine the accuracy of ERnet, we generated synthetic ground truth data on which the performance of the method could be tested. First, we generated data to test semantic segmentation performance. To do this, we used real SIM data of ER networks on which we applied the well-established Trainable Weka Segmentation machine learning algorithm[15]. This produced ground truth images for which ER structures were classified into tubules, sheets and SBTs (Fig. 3a).

The same SIM images on which the above Weka approach was used to generate ground truth data were then processed by ERnet, and the results were compared pixel by pixel. The ground truth test demonstrated a pixel accuracy for ERnet segmentation of between 92% and 99% compared to that for the ground truth data (Fig. 3c). In another test, we used the segmented images obtained with the Trainable Weka algorithm and fed this as input to ERnet. In this case again, the result was nearly identical to the ground truth.

https://doi.org/10.1038/s41592-023-01815-0

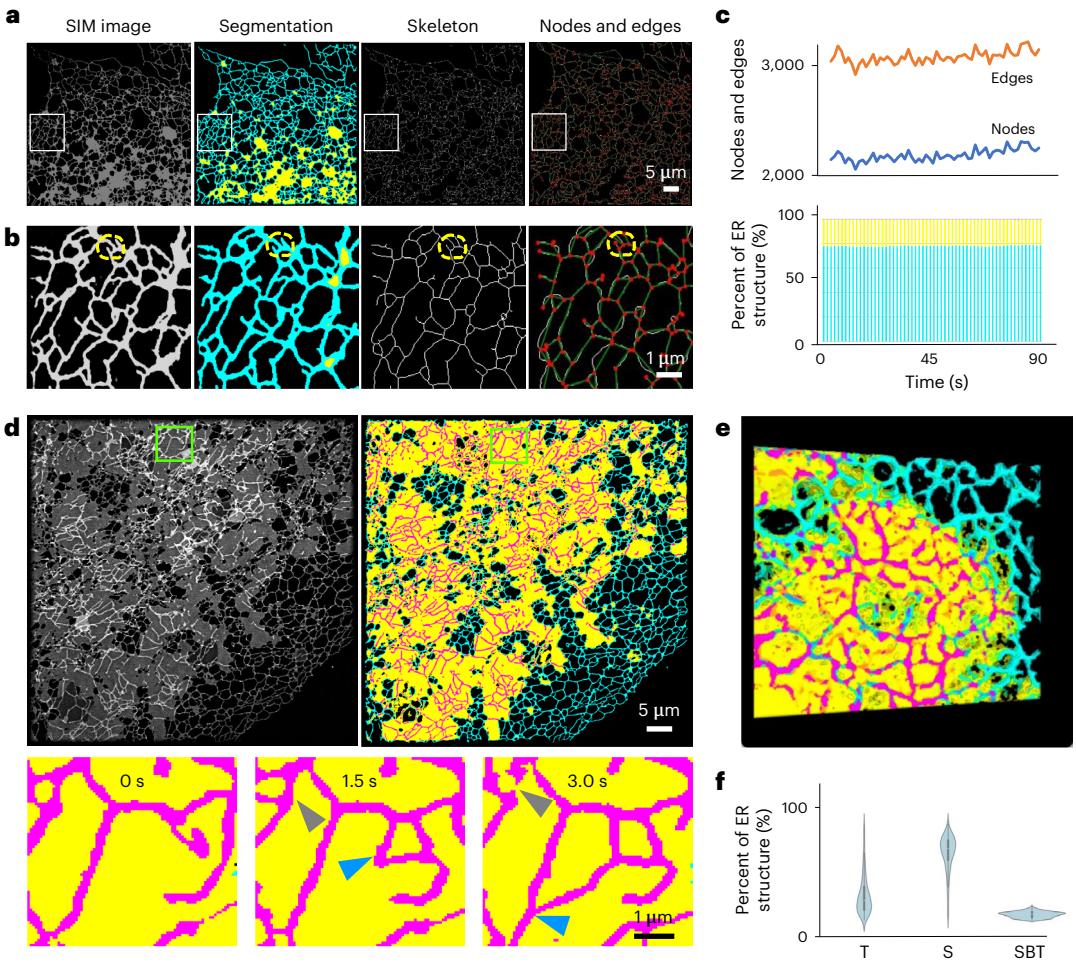


Fig. 2 | Semantic segmentation of the ER and classification of tubules and sheets. a, An example of a segmentation result from video-rate SIM images of the ER. Left to right, (1) SIM image, (2) segmentation of ER tubular (cyan) and sheet (yellow) regions, (3) skeletonization of the tubular domain and (4) identification of nodes (red spots) and edges (green lines) based on the skeleton structure. Scale bar, 5 µm. **b**, Magnified regions of **a**. Yellow dashed circles indicate nodes that are closely positioned but can still be identified by ERnet. Scale bar, 2 µm. **c**, Quantitative analysis of the ER shown in **a**. Top, quantification of edges and nodes of the ER tubules of time-lapse frames. Bottom, percentage of the ER tubules (cyan) and sheets (yellow) of time-lapse frames (1.5 s per frame) (Source Data Fig. 2c). **d**, A representative frame from time-lapse images shows the structure of SBTs (1.5 s per frame). Top left, a SIM image of the ER structure. Top right, segmentation of the three ER structures: SBTs (magenta), sheets (yellow) and tubules (cyan). Bottom, three sequential frames showing the dynamic reshaping of SBTs from the above green boxed region. Blue arrows indicate a continuously elongating SBT, and gray arrows indicate a retraction. Scale bars, 5 µm (top) and 2 µm (bottom). See Source Data Fig. 2d for quantitative analysis. **e**, Volumetric view of a 3D reconstruction of the sectioning SIM showing that the SBTs (magenta) are embedded in sheet domains (yellow). Scale bar, 2 µm (bottom). **f**, Violin plots of the percentages of tubules (T), sheets (S) and SBTs in COS-7 cells ($N = 500$), showing that the presence of SBTs is a common feature of the ER network. In the violin plots, the white dot represents the median value of the data, the thick bar represents the interquartile range, and the thin bar represents the rest of the data distribution (Source Data Fig. 2f).

In addition to the ground truth test for semantic segmentation, we also tested the accuracy of the connectivity analyses. To do this, we generated ground truth data of ER tubular domains by creating synthetic ER skeletons. We then widened and blurred the skeletons and added image noise to mimic ER structures recorded with optical microscopy (for details on this process, see Extended Data Fig. 4a and its descriptive caption). After this, the synthetic images were processed by ERnet to identify nodes and edges and to derive metrics for ER connectivity (Fig. 3d,e). ERnet reached accuracies ranging from 96% to 99% for identification of nodes and edges (Fig. 3f). Even in dense regions of the tubular network (magnified regions, Fig. 3e), ERnet still achieved a high precision to capture nodes and edges, with little difference found between the ERnet result and the ground truth data. Additionally, we quantified the differences in connectivity metrics obtained from ERnet and ground truth data (Fig. 3f). Because the assortativity metric ranges over very small scales, for example, from −0.05 to 0.08, even minor changes in connectivity can lead to large fluctuations of the former. Nevertheless, observed changes in metrics are still significantly smaller than those associated with the varying phenotypes reported in the following context.

Next, we tested the performance of ERnet on ground truth images in which we added variable levels of noise (Extended Data Fig. 4). The purpose was to provide a metric with which a user can decide upfront whether a given dataset obtained on a microscope is of sufficient quality to trust the ERnet output. We found that ERnet produced repeatedly reliable outputs for both connectivity and topology features for image data featuring SNRs better than ~5 (Extended Data Fig. 4). By analyzing the SNR obtained with a given experimental setup, users can objectively assess the quality of segmentation results, irrespective of where and how the data were obtained.

ERnet performs on various cell types and imaging modalities

To demonstrate the versatility and robustness of ERnet in different research scenarios, we validated the method on a range of datasets

https://doi.org/10.1038/s41592-023-01815-0

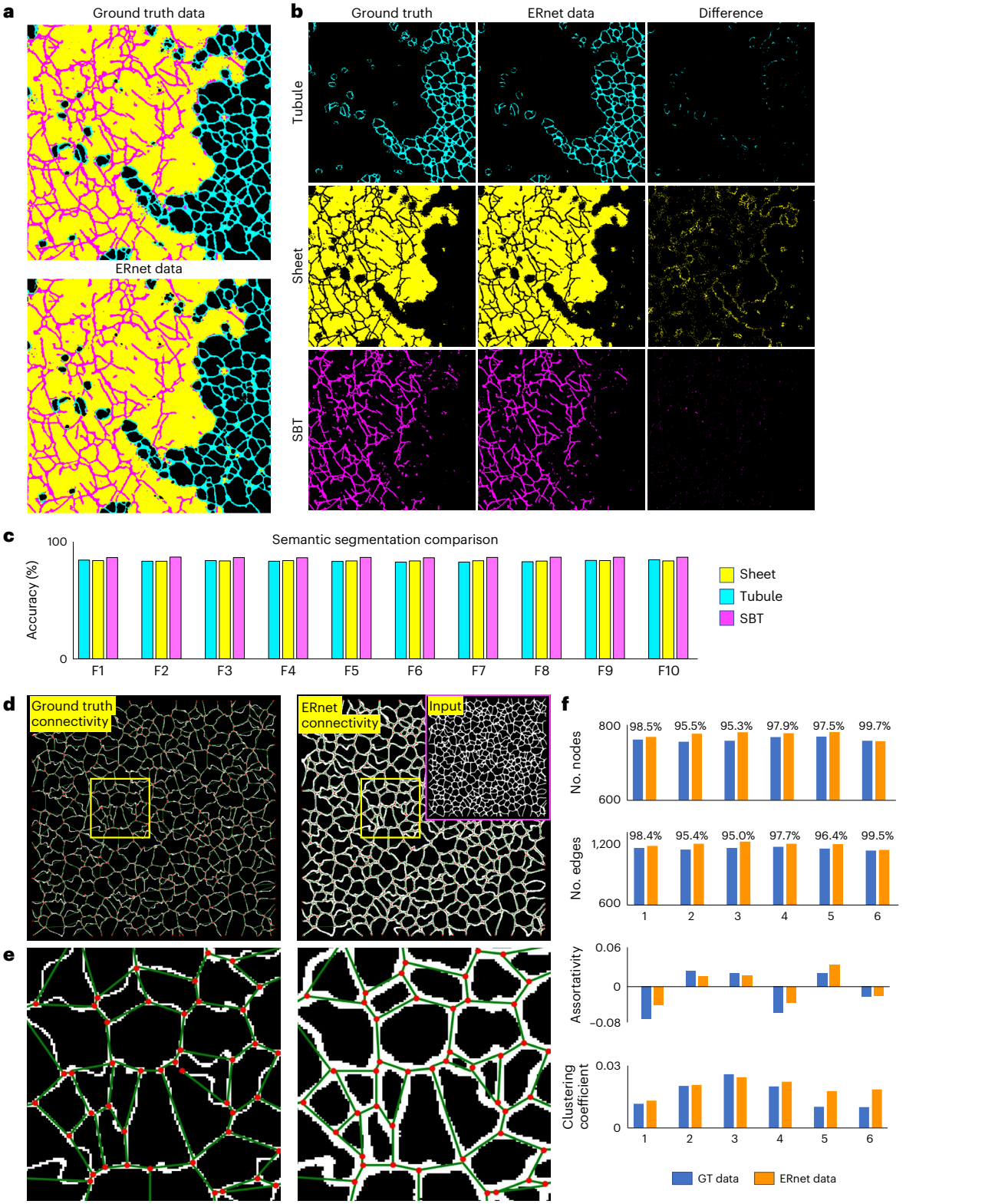


Fig. 3 | Ground truth test for segmentation and connectivity. a, Comparison between ground truth data and ERnet data for ER tubular (cyan), sheet (yellow) and SBT (magenta) domains. Comparisons were repeated three times independently with similar results shown in **c**. **b**, Comparison of each channel for the image data above. **c**, Quantification of pixel differences from the three image channels. F1–F10 are frames 1–10, respectively, in sequentially recorded ER images (Source Data Fig. 3c). **d**, Comparison of ground truth data (synthetic ER tubular network) and ERnet results. The top right inset and that framed in a magenta box present the whole field of view of the ground truth data, which were input into ERnet. **e**, A magnified region of the highlighted sections in **a** showing that the connectivity revealed by ERnet is nearly identical to the ground truth. Red spots, nodes; green lines, edges. **f**, Comparison of the connectivity metrics. GT data, ground truth data. Numbers on the *x* axis indicate the image sample number (Source Data Fig. 3f).

obtained by us and others. Fig. 4 presents the analysis of images obtained using different microscopy techniques including widefield, confocal and Airyscan microscopy. Even though ERnet's precision may depend on the spatial resolution of the corresponding images, it performed well for all imaging techniques, with all tubules and sheets clearly classified and quantified (Source Data Fig. 4). Furthermore, we also performed validation tests for varying cell types commonly used in cell biology research, such as HEK, CHO and SH-SY5Y cells, and primary cultures of hippocampal neurons and glial cells derived from embryonic rats (Source Data Fig. 4). Further data from plant cells[13] and publicly available datasets[23,24] published by other authors using different experimental setups are shown in Extended Data Fig. 5. Although the specific ER phenotypes varied among the cell types, ERnet was able to robustly identify the corresponding tubular and sheet domains and performed subsequent quantitative analyses following segmentation. The model did not have to retrained for any of these scenarios, and no preprocessing of the raw data was necessary before segmentation by ERnet, demonstrating the generality of the model and its ease of application.

ERnet provides detailed connectivity data on ER networks

ERnet can be used to quantify the connectivity of edges and nodes before plotting a corresponding connectivity graph (Fig. 5a). The connectivity graph highlights that the network of the ER is largely composed of three-way junctions (red nodes, Fig. 5a), while ER edges are capped with growth ends (green nodes, Fig. 5a).

To assess the integrity of the ER, we defined each disconnected ER region as a fragment. Although the number of fragments during ER reshaping fluctuates (Fig. 5b), ERnet revealed that, in a typical healthy cell, the majority of all edges and nodes are contained in a single large fragment at all times (over 92% of all 3,000 nodes and 95% of all 4,000 edges in the example shown). As quantitative parameters, we defined node- and edge-assembly ratios (the number of nodes or edges in the largest fragment divided by the total number of nodes or edges, respectively) (Fig. 5c). According to the definition, these values range from close to 0 (fully fragmented ER) to 1 (fully connected). Additionally, ERnet quantified the degrees of the ER nodes, that is, how many edges (tubules) connect to each node (junction). As shown in Fig. 5d, three-way junctions were the most abundant and represented 66% of all junction types in this example. Despite the prevailing model of ER morphology, in which three-way junctions interconnect to form the whole ER tubular network, ERnet also identified nodes connected with more than three edges (tubules), that is, multi-way junctions. The presence of multi-way junctions indicates the heterogeneous connectivity of ER tubules that are organized in a higher order of complexity than previously assumed.

Next, the assortativity and clustering coefficients (Fig. 5e,f), which describe connectivity patterns of nodes, were calculated based on the above metrics. The assortativity coefficient measures the tendency of nodes to connect with others of the same degree[25]. In a network with a high assortativity coefficient, most nodes are connected in a similar way with their neighbors (for example, via three-way junctions). The clustering coefficient, on the other hand, reflects the distribution of nodes within the whole network (for example, clusters of multi-way junctions may be separated from other clusters by junctions of lower degree). For a graphical explanation of these concepts, the reader is referred to Extended Data Fig. 6.

Assortativity coefficients range from −1 (network fully heterogeneous in connectivity, that is, nodes only connect with those of different degrees) to +1 (network fully homogeneous in connectivity, that is, nodes only connect with those of same degree). Similarly, for clustering coefficients, 1 describes a network in which all the nodes and edges are clustered, while 0 refers to no clustering. Fig. 5e shows the ER as a weak assortative network, which suggests a tendency, albeit a weak one, of nodes to connect with nodes of the same degree. In Fig. 5f, we show how the degree of clustering can change over time in an ER network. Tubules and junctions reorganize themselves rapidly, both within localized and global domains. Frequent events include the merging of multiple tubules forming clusters of nodes, but these then disassemble transiently. Overall, the data indicate that the network features a high degree of structural homogeneity, and local clustering is not a dominating feature to affect the overall phenotype.

To further investigate the structural dynamics of the ER, we tracked the lifetime of multi-way junctions and their transitions from multi-way junctions to three-way junctions. Fig. 5g,h shows rapid transitions between three-way (yellow arrows) and multi-way (blue arrows) junctions driven by ER tubule reshaping. As shown in these cases, the formation of four-way or five-way junctions needs simultaneous connections of more than three tubules at the same junction, which occurs with a lower chance than the formation of a three-way junction, which only requires the connection of three tubules. Additionally, any movement of a tubule away from its multi-way junction can lead to the collapse of this junction and the generation of at least two three-way junctions. Therefore, as shown in Fig. 5i, the average lifetime of a multi-way junction is much shorter, that is, less than a third (9.0 s versus 30.8 s) of that of a three-way junction.

We also examined whether our 2D network approach is valid to segment ER structures, which are 3D in nature. We performed two different tests with both COS-7 and U2OS cells, which are the canonical models in fluorescence microscopy-based studies of the ER and for which ERnet was developed. For these flat cell types, we could confirm that a 2D analysis was sufficient to represent the ER network topology (Extended Data Fig. 7).

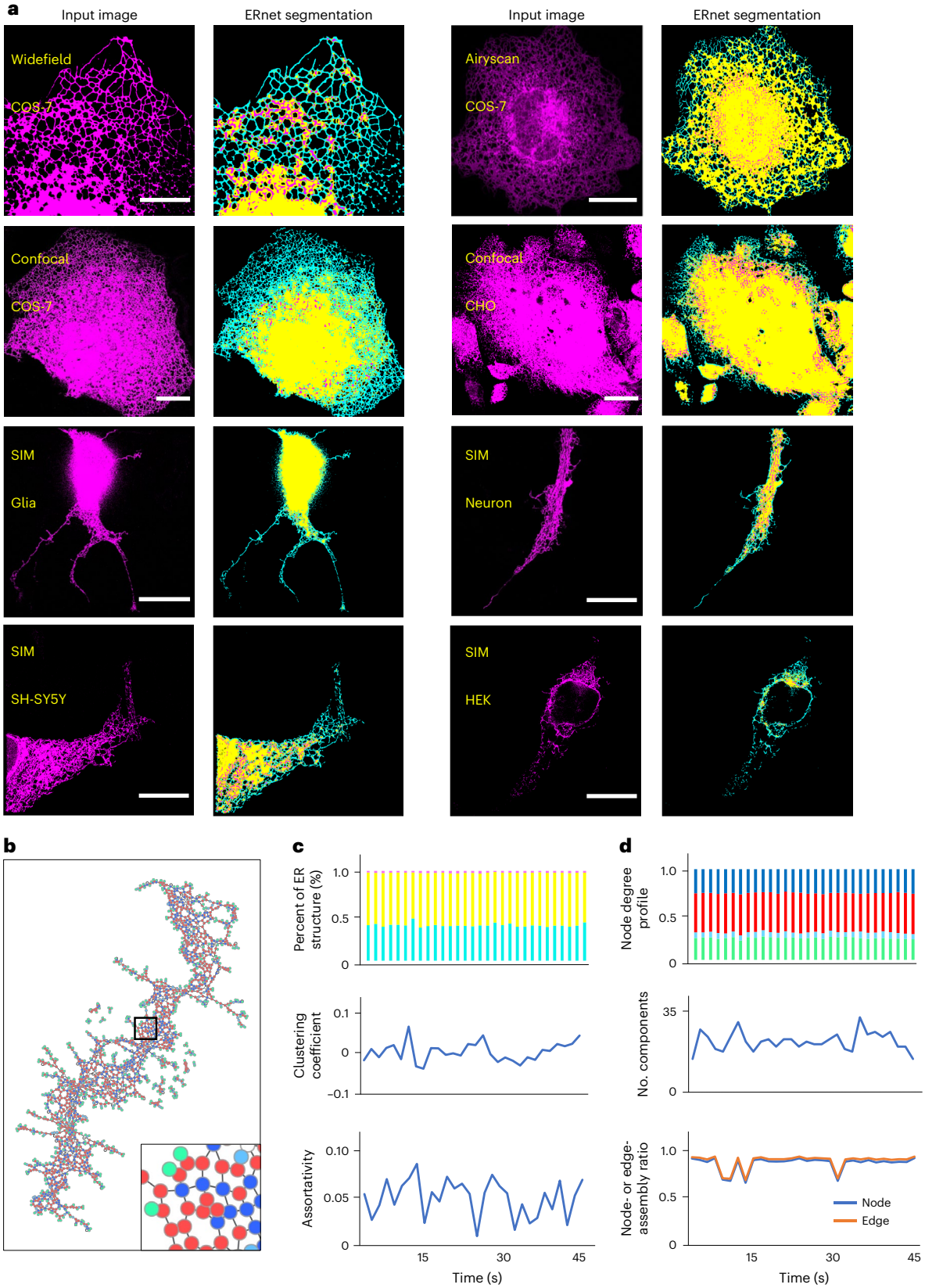
ERnet can characterize complex ER phenotypes

ER morphological defects caused by mutations in genes encoding ER-reshaping proteins or by metabolic perturbations have been linked to a variety of human diseases[1,2,4]. However, the exact phenotypical ER disruption under these conditions has not yet been sufficiently characterized. Using ERnet, we first analyzed ER morphological defects in stress models mimicking the ER phenotypes in two neurodegenerative diseases, namely, HSP and NPC. The inherited neurological disorder HSP can be characterized by progressive lower-limb weakness and muscle stiffness, which are caused by mutations in genes encoding ER-reshaping proteins such as atlastin (ATL)[26] and protrudin[27]. We used ERnet to examine ER morphology defects in individual cells of different

Fig. 4 | Robust performance of ERnet in a versatility test. a, A variety of cell lines with different ER morphologies were imaged by different microscopy techniques to investigate the robustness and versatility of ERnet. ER structures of COS-7, HEK, CHO and SH-SY5Y cells and primary cultures of hippocampal neurons and glial cells were tested as well as images acquired by widefield, confocal and Airyscan microscopy (1.5 s per frame). Scale bars, 20 µm. **b**, The topology of an ER tubular network of the COS-7 cell from the confocal image shown in a is represented by a connectivity graph. Nodes of different degrees are labeled with different colors: green (degree 1), light blue (degree 2), red (degree 3) and dark blue (degree >3). Bottom right, a magnified region of the black boxed part in the connectivity graph, demonstrating complex connectivity revealed by ERnet from the confocal microscopy image. The following analysis of **c**,**d** is based on this image data. **c**, Quantitative analysis of the ER structure of the image data from **b** reveals topology features of the ER tubular network. Top, percentage of ER tubules (cyan), sheets (yellow) and SBTs (magenta) of the time-lapse frames (43.5 s at 1.5 s per frame). Middle and bottom, changes of assortativity and clustering coefficients in time-lapse images (Source Data Fig. 4c,d). **d**, Quantitative analysis of the connectivity of the ER tubular network in the cell from **b**. Top, quantification of the nodes of different degrees, showing a dominance of third-degree nodes (three-way junctions). Middle, number of components (ER fragments) in time-lapse images. Bottom, changes in the node or edge ratio over time.

https://doi.org/10.1038/s41592-023-01815-0

models by measuring two topological features, that is, the degree of ER tubule fragmentation and the heterogeneity in these tubular connections. Compared with control cells, ATL (ATL2 and ATL3) KO[28] leads to a collapse of ER network integrity. Such ER fragmentation was clearly revealed in ATL KO cells by the increasing number of fragments and a 20-fold reduction of the node-assembly ratio (90% in control versus 4.5% in ATL KO) (Fig. 6a and Supplementary Videos 5 and 6). ERnet also highlighted that the lack of ATL significantly altered connectivity in the ER tubular network, as witnessed by a reduced percentage of three-way junctions among all nodes (22% versus 65% in control) and by the disorganized connectivity (−0.25 in assortativity). These measurements provide quantitative rather than descriptive evidence



https://doi.org/10.1038/s41592-023-01815-0

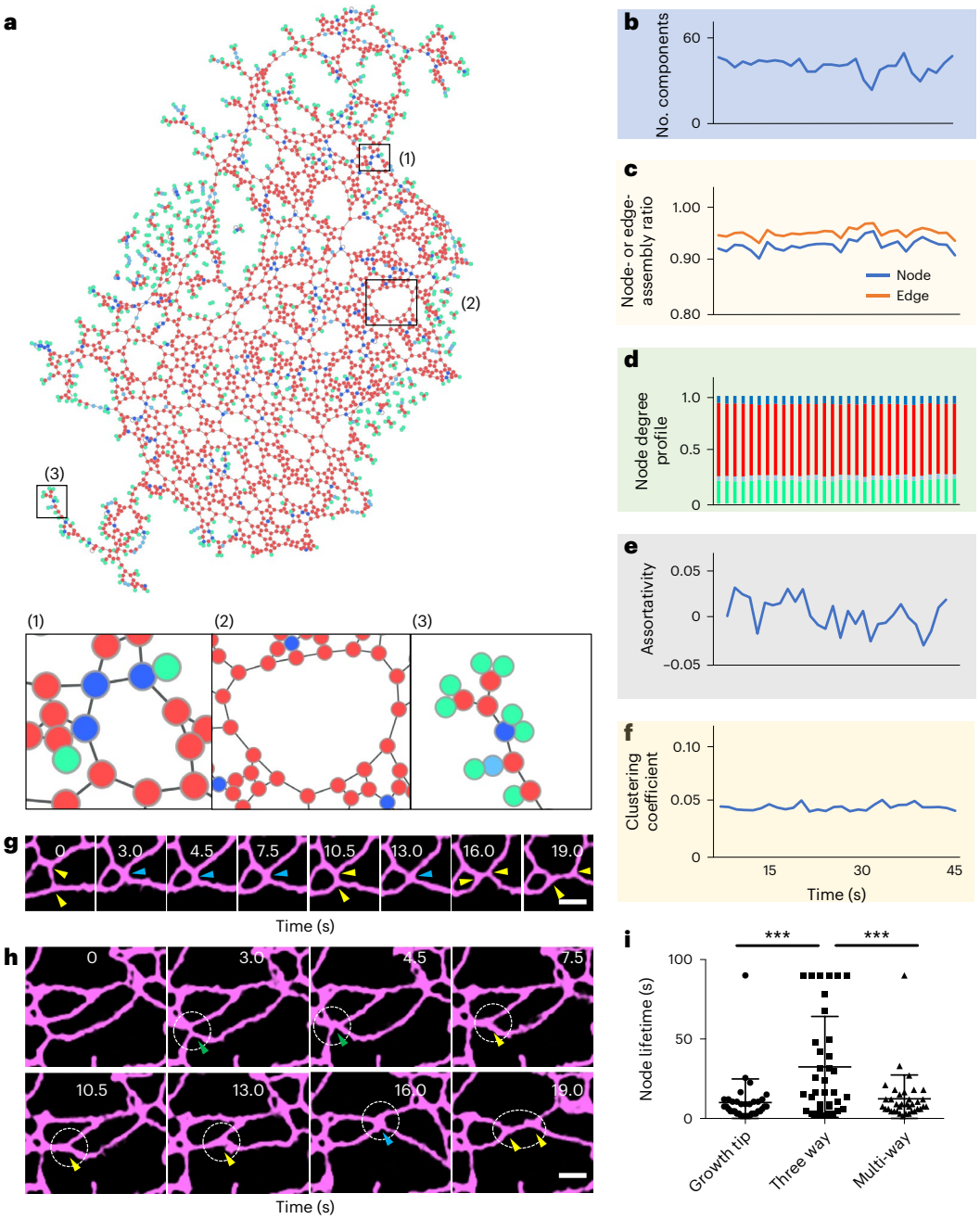


Fig. 5 | Quantitative analysis by ERnet reveals complex connectivity of the ER tubular network. a, The topology of an ER tubular network is represented by a connectivity graph. (1), a representative region of multi-way junctions (dark blue spots); (2), a polygonal structure organized by three-way junctions and tubules; (3), a representative region of ER tubular growth tips (green spots). **b**–**f**, Quantitative analysis of the cell shown in **a** over a time window of 45 s at 1.5 s per frame (Source Data Fig. 5f). **b**, Number of components (ER fragments) in time-lapse images. **c**, Changes of the node- or edge-assembly ratio over time. **d**, Quantification of nodes of different degrees, showing a dominance of third-degree nodes (three-way junctions). Same color scheme as that in **a**. **e**,**f**, Changes in assortativity and clustering coefficients in time-lapse images. **g**,**h**, Examples of transitions between three-way (yellow arrows) and multi-way (yellow arrows, three-way; blue arrows, four-way; green arrows, five-way) junctions. Scale bar, 1 µm. **i**, Quantification of the lifetime of junctions (nodes) with different degrees. Data are presented as mean ± s.e.m., ***$P < 0.001$, Tukey's one-way ANOVA. $N = 12$ events per condition per experiment from three independent experiments and 36 events per condition were analyzed in total. P values, growth tip versus multi-way junctions, 0.8947; growth tip versus three-way junctions, 0.0001; three-way versus multi-way junctions, 0.0006 (Source Data Fig. 5i).

of ATL's role in ER tubular network formation, which has previously been reported to be crucial for the fusion of ER membranes and thus the formation of continuous networks[26]. With these quantitative analyses, we can compare morphological defects caused by different treatments. In another model of HSPs, depletion of protrudin (Extended Data Fig. 8) also resulted in ER tubular network fragmentation (350 fragments) (Supplementary Video 7) and in disorganized connectivity, however, to a lesser extent. A further metric suitable for the comparison of ER health under different treatments is the size of the ER, which is revealed by the connectivity graph. An ATL KO cell that was more fragmented than a protrudin-knockdown cell suffered from more severe shrinkage of the ER with a smaller number of nodes and edges (Fig. 6a), indicating that ER membranes may be degraded or recycled in response to stresses.

https://doi.org/10.1038/s41592-023-01815-0

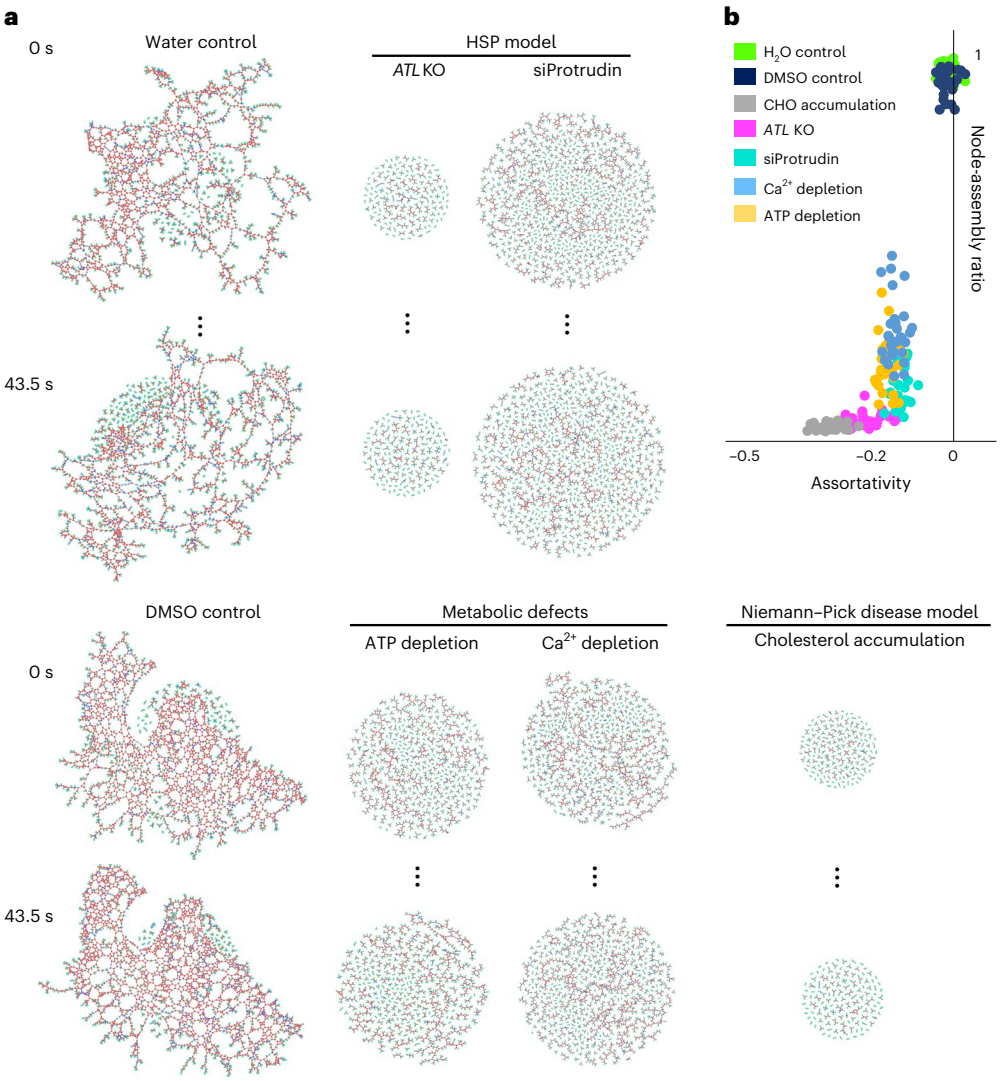


Fig. 6 | Quantitative analysis of ER phenotypic characteristics in disease-associated models. a, Connectivity graphs of ER structures in models mimicking phenotypes of HSPs and NPC and metabolic stress induced by calcium and ATP depletion. Nodes of different degrees are labeled with different colors: green (degree 1), light blue (degree 2), red (degree 3) and dark blue (degree >3). Note that the graphs represent data from the whole field of view imaged in the microscope. Connectivity graphs are symbolic reductions of ER networks for easy visualization of network topology. Graphs should not be mistaken for actual spatial representations of ER networks. Highly connected networks (controls, left) appear more amorphous than strongly fragmented networks (stressed cells, middle). Raw image data are presented in Supplementary Videos 5–11 (1.5 s per frame). siProtrudin, protrudin knockdown. **b**, Topological features of the ER tubular network in the conditions in **a** were quantitatively analyzed by ERnet. The effects on ER structures from different treatments can be directly visualized and compared by plotting the distribution of the node-assembly ratio (*y* axis) and assortativity coefficient (*x* axis). The analysis of ER phenotype, such as that in ATL KO cells, demonstrated severe fragmentation and altered connectivity in the numerical data plot (Source Data Fig. 6b). DMSO, dimethyl sulfoxide.

Next, we induced cholesterol accumulation in lysosomes by U18666A administration to the cell, which induces blockage of cholesterol transfer from lysosomes to the ER in NPC[29]. Accumulation of cholesterol in lysosomes leads to lysosome deposition in perinuclear regions and, therefore, affects ER structure and distribution[3]. However, the exact ER morphological defects have not yet been characterized. ERnet revealed that the ER of U18666A-treated cells features a disassortative network (−0.34), and its low node-assembly ratio (2.6%) suggests a highly fragmented structure (Fig. 6a,b and Supplementary Video 8), which highlights the fact that lysosomal defects can strongly affect the ER and thus provides for a useful tool with which to improve an understanding of organelle dysfunction in NPC.

Finally, we tested the performance of ERnet in cells upon ER collapse under metabolic manipulations that substantially affect overall homeostasis inside the cell. A SIM video showed that the ER largely loses its dynamic reshaping capabilities upon administration of the store-operated calcium entry inhibitor SKF96365 (ref. 30) (Supplementary Video 9). In the connectivity graph, the ER became largely fragmented and featured as a disassortative network (Fig. 6a,b). Compared with SKF96365, NaN_3 depletes ATP[31], thus capping support for all energy-consuming processes inside the cell, including ER tubule elongation, retraction and membrane fusion. Therefore, ATP depletion by NaN_3 was expected to substantially affect the structural dynamics of the ER. ERnet revealed the level of ER network fragmentation resulting from a lack of ATP (Fig. 6a,b and Supplementary Video 10); however, the phenotypes were not equivalent to those observed upon depletion of ER-reshaping proteins: for example, the node-assembly ratio in ATP-depleted cells was found to be nearly fourfold of that in ATL KO cells (0.19 versus 0.05).

Overall, while ERnet provides a quantitative assessment of overall network topology, it is also sensitive enough to detect subtle

changes in local ER morphology, valuable attributes in the investigation and differentiation of ER-related phenotypes of disease.

Discussion

A measurement of cellular organelle properties such as shape, position and mobility provides a quantitative basis for analyzing the structure and function of organelles in both fundamental and therapeutic research. Here, we introduce ERnet, a versatile tool that performs robust and precise segmentations and analyses of ER structures under a variety of conditions.

The accuracy of ERnet's semantic segmentation algorithm is a result of its model design. In contrast to state-of-the-art CNN models commonly used for image segmentation, ERnet is constructed in a Vision Transformer architecture that outperforms CNNs in terms of image-classification accuracy and requires much smaller computational resources[19,32]. Another advantage of our design is a capability for temporal domain analyses of objects in sequenced image data. We integrated two attention mechanisms, multi-head self-attention (MSA)[33] and channel attention[34], into the Transformer architecture. These mechanisms greatly enhance the learning ability of ERnet in classifying ER structures in the spatio-temporal domain. While machine learning methods have previously been implemented for denoising images of ER structures[35], reconstructing ER structures based on electron microscopy images[36] and identification ER stress marker whorls[37], ERnet is capable of video-rate image segmentation and analysis of live cells, further extending the deep learning toolbox for biomedical research.

Through application of ERnet, we were able to characterize and quantify structural features of dynamic ER networks. First, we found that the dominance of three-way junctions is a necessity to produce a continuous ER network that can spread throughout the cell. While we observed a prevalence of three-way junctions, we found that 20% of healthy ER furthermore consists of multi-way junctions (degree >3). By contrast, all stress manipulations of ER morphology, including models of HSPs and NPC, resulted in the fragmentation of ER structures to varying extents (Fig. 6).

Nixon-Abel et al.[5] and Pain et al.[13] also made use of microscopy data in the analysis of ER dynamics. However, their work focused on a very different set of ER phenomena than what we present here. Nixon-Abel et al.[5] analyzed transient dynamics of individual tubules (for example, lateral tubule oscillations). Pain and colleagues designed AnalyzER to extract metrics of plant ER tubules, such as width, length and cross-sectional area. By contrast, our work focuses on global network topology and integrity, which are key features associated with physiological and stress states. Our aim is to provide a robust and powerful tool for investigation of therapeutic strategies against ER-associated disease. Apart from this, ERnet, driven by deep learning to classify ER structures, can identify subtle changes in the whole ER and display the difference in quantitative plots. The connectivity graph is a unique feature of ERnet. It is a visual tool to display the connected parts of the ER and provides for a rapid visual cue on the degree of network integrity.

An advantage of the use of deep learning in biological imaging is that it facilitates discovery of new biological phenomena. The sensitivity of ERnet to changing structural features led to the identification of SBTs. These ER components share similar structures and dynamics with the tubules that radiate from the sheet domains toward the periphery of the cell; however, their existence in the sheet domain greatly extends the coverage of the tubular ER toward the cell center and even close to the nucleus. We note that SBTs are evident also in data presented in previous reports, such as that by Schroeder and colleagues[38] (for example, Figs. 1e,h, 3b and 4a), but the phenomenon was not recognized specifically. In our method, SBTs are classified in addition to sheets and tubules on their own. While ERnet can be used with any imaging technique, conventional or super-resolved, its ability to detect and classify SBTs does depend on SNR and image resolution. Therefore, some differences are expected in output produced from very different imaging methods. We also note that ER topology can vary substantially from cell to cell and do not recommend conclusions to be drawn from data that are not representative of the whole cell population. How SBTs are regulated in both physiological and pathological conditions will be an important question for future studies.

Similar to all segmentation and classification methods, including those performed by humans, ERnet is necessarily limited by the quality of the input data. We found that, for SNRs greater than around 5, ERnet reliably quantifies topology structures for any microscopy method of appropriate image resolution. Because we optimized ERnet for high-throughput analysis, the algorithm treats ER networks as 2D structures for computational efficiency. While we saw no problems with this for the cell types that we analyzed, one needs to be cautious when applying the method to highly 3D networks. ERnet could be extended to three dimensions and integrated with further organelle-analysis tools, for example, methods for characterization of lysosomes and mitochondria, to permit comprehensive investigations of organelle–organelle interactions and their role in the development, aging and degeneration of cells.

We believe that our work demonstrates an efficient tool for precise structure segmentation and multi-parameter analysis of ER phenotypes. Its user-friendly graphical interface and automatic batch-processing capabilities obviate the need for manual annotation.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-023-01815-0.

References

1. Schwarz, D. S. & Blower, M. D. The endoplasmic reticulum: structure, function and response to cellular signaling. *Cell. Mol. Life Sci.* **73**, 79–94 (2016).
2. Schönthal, A. H. Endoplasmic reticulum stress: its role in disease and novel prospects for therapy. *Scientifica* **2012**, 857516 (2012).
3. Lu, M. et al. The structure and global distribution of the endoplasmic reticulum network are actively regulated by lysosomes. *Sci. Adv.* **6**, eabc7209 (2020).
4. Obara, C. J., Moore, A. S. & Lippincott-Schwartz, J. Structural diversity within the endoplasmic reticulum—from the microscale to the nanoscale. *Cold Spring Harb. Perspect. Biol.* **19**, a041259 (2022).
5. Nixon-Abell, J. et al. Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. *Science* **354**, aaf3928 (2016).
6. Guo, Y. et al. Visualizing intracellular organelle and cytoskeletal interactions at nanoscale resolution on millisecond timescales. *Cell* **175**, 1430–1442 (2018).
7. Zhao, W. et al. Sparse deconvolution improves the resolution of live-cell super-resolution fluorescence microscopy. *Nat. Biotechnol.* **40**, 606–617 (2022).
8. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
9. Fischer, C. A. et al. MitoSegNet: easy-to-use deep learning segmentation for analyzing mitochondrial morphology. *iScience* **23**, 101601 (2020).
10. Lefebvre, A. E., Ma, D., Kessenbrock, K., Lawson, D. A. & Digman, M. A. Automated segmentation and tracking of mitochondria in live-cell time-lapse images. *Nat. Methods* **18**, 1091–1102 (2021).
11. Hollandi, R. et al. nucleAIzer: a parameter-free deep learning framework for nucleus segmentation using image style transfer. *Cell Syst.* **10**, 453–458 (2020).

12. English, A. R. & Voeltz, G. K. Endoplasmic reticulum structure and interconnections with other organelles. *Cold Spring Harb. Perspect. Biol.* **5**, a013227 (2013).
13. Pain, C., Kriechbaumer, V., Kittelmann, M., Hawes, C. & Fricker, M. Quantitative analysis of plant ER architecture and dynamics. *Nat. Commun.* **10**, 984 (2019).
14. Garcia-Pardo, M. E., Simpson, J. C. & O'Sullivan, N. C. A novel automated image analysis pipeline for quantifying morphological changes to the endoplasmic reticulum in cultured human cells. *BMC Bioinformatics* **22**, 427 (2021).
15. Arganda-Carreras, I. et al. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33**, 2424–2426 (2017).
16. Belthangady, C. & Royer, L. A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat. Methods* **16**, 1215–1225 (2019).
17. Lee, T. C., Kashyap, R. L. & Chu, C. N. Building skeleton models via 3-D medial surface axis thinning algorithms. *CVGIP Graph. Models Image Process.* **56**, 462–478 (1994).
18. Peixoto, P. T. The graph-tool Python library (figshare, 2014).
19. Dosovitskiy, A. et al. An image is worth 16×16 words: transformers for image recognition at scale. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2010.11929 (2020).
20. Liu, Z. et al. Swin Transformer V2: scaling up capacity and resolution. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition* 12009–12019 (IEEE/CVF, 2022).
21. Mateus, D., Marini, E. S., Progida, C. & Bakke, O. Rab7a modulates ER stress and ER morphology. *Biochim. Biophys. Acta Mol. Cell Res.* **1865**, 781–793 (2018).
22. Chang, J., Lee, S. & Blackstone, C. Protrudin binds atlastins and endoplasmic reticulum-shaping proteins and regulates network formation. *Proc. Natl Acad. Sci. USA* **110**, 14954–14959 (2013).
23. Qin, J. et al. ER–mitochondria contacts promote mtDNA nucleoids active transportation via mitochondrial dynamic tubulation. *Nat. Commun.* **11**, 4471 (2020).
24. Qiao, C. et al. Evaluation and development of deep neural networks for image super-resolution in optical microscopy. *Nat. Methods* **18**, 194–202 (2021).
25. Newman, M. E. J. Assortative mixing in networks. *Phys. Rev. Lett.* **89**, 208701 (2002).
26. Zhao, X. et al. Mutations in a newly identified GTPase gene cause autosomal dominant hereditary spastic paraplegia. *Nat. Genet.* **29**, 326–331 (2001).
27. Mannan, A. U. et al. ZFYVE27 (SPG33), a novel spastin-binding protein, is mutated in hereditary spastic paraplegia. *Am. J. Hum. Genet.* **79**, 351–357 (2006).
28. Sun, S. et al. Identification of endoplasmic reticulum-shaping proteins in Plasmodium parasites. *Protein Cell* **7**, 615–620 (2016).
29. Ko, D. C., Gordon, M. D., Jin, J. Y. & Scott, M. P. Dynamic movements of organelles containing Niemann–Pick C1 protein: NPC1 involvement in late endocytic events. *Mol. Biol. Cell* **12**, 601–614 (2001).
30. Merritt, J. E. et al. SK&F 96365, a novel inhibitor of receptor-mediated calcium entry. *Biochem. J.* **271**, 515–522 (1990).
31. McAbee, D. D. & Weigel, P. H. ATP depletion causes a reversible redistribution and inactivation of a subpopulation of galactosyl receptors in isolated rat hepatocytes. *J. Biol. Chem.* **262**, 1942–1945 (1987).
32. Paul, S., & Chen, P. Y. Vision transformers are robust learners. In *Proc. AAAI Conference on Artificial Intelligence* Vol. 36, 2071–2081 (AAAI, 2022).
33. Vaswani, A. et al. Attention is all you need. *Advances in Neural Information Processing Systems* 30 (NIPS) (2017).
34. Christensen, C. N., Lu, M., Ward, E. N., Lio, P. & Kaminski, C. F. Spatio-temporal Vision Transformer for super-resolution microscopy. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2203.00030 (2022).
35. Qiao, C. et al. Rationalized deep learning super-resolution microscopy for sustained live imaging of rapid subcellular processes. *Nat. Biotechnol.* **41**, 367–377 (2023).
36. Liu, J. et al. Automatic reconstruction of mitochondria and endoplasmic reticulum in electron microscopy volumes by deep learning. *Front. Neurosci.* **14**, 599 (2020).
37. Guo, Y. et al. Deep learning-based morphological classification of endoplasmic reticulum under stress. *Front. Cell Dev. Biol.* **9**, 767866 (2022).
38. Schroeder, L. K. et al. Dynamic nanoscale morphology of the ER surveyed by STED microscopy. *J. Cell Biol.* **218**, 83–96 (2019).

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s) under exclusive license to Springer Nature America, Inc. 2023

Methods

Cell culture

COS-7 cells were purchased from the American Type Culture Collection (CRL-1651, ATCC). COS-7 cells were grown in T75 or T25 flasks or six-well plates by incubation at 37 °C in a 5% CO_2 atmosphere. Complete medium for normal cell growth consisted of 90% DMEM, 10% FBS and 1% streptomycin. Cells were kept in the logarithmic phase of growth and passaged on reaching 70–80% confluence (approximately every 3–4 d). Medium was changed every 2 or 3 d. For SIM imaging experiments, COS-7 cells were plated onto Nunc Lab-Tek II Chambered Coverglass (Thermo Fisher Scientific, 12-565-335) to achieve ~70% confluence before transfection.

COS-7 cells were transfected with mEmerald-Sec61b-C1 (Addgene, 90992, gifted by J. Lippincott-Schwartz, Janelia Research Campus) as indicated with Lipofectamine 2000 according to the manufacturer's protocol 24–48 h before imaging. Cells were stained with SiR-lysosome at 1 µM for 4 h before imaging. Cells were imaged in a microscope stage-top micro-incubator (OKO Lab) with continuous air supply (37 °C and 5% CO_2). Cells were treated with U18666A (662015, Sigma) at 10 µM for 24 h to block cholesterol transfer from lysosomes to the ER before imaging. Cells were treated with SKF96365 (S7809, Sigma) at 100 µM for 3 h to deplete calcium before imaging. Cells were treated with NaN_3 (0.05%, wt/vol) and 2-deoxyglucose (20 mM) for 2 h to deplete ATP before imaging. SH-SY5Y cells (CRL-2266, ATCC) were cultured and imaged as previously described[39]. The ATL KO model[28] was constructed by deleting ATL2 and ATL3 using the CRISPR–Cas9 system in COS-7 cells (ATL1 is not detectable in COS-7 cells), a gift from J. Hu, Chinese Academy of Sciences, China. CHO-K1 (CCL-61, ATCC) cells were purchased from ATCC and were cultured in Ham's F-12 Nutrient Mixture medium supplemented with 10% FBS, 2 mM L-glutamine and 100 U ml^{-1} penicillin–streptomycin. Cells were transfected with pFLAG_ER mCherry[40]. U2OS cells (HTB-96, ATCC) were cultured in DMEM supplemented with 10% FBS, 2 mM L-glutamine and 100 U ml^{-1} penicillin–streptomycin. Cells were transfected with pFLAG_ER mCherry. Primary tissues, including hippocampal neurons and glial cells, were isolated from rats at postnatal day 1 (Sprague–Dawley rats from Charles River) and cultured as described before[41]. HEK293T cells (CRL-3216, ATCC) were cultured and imaged as described before[42].

Small interfering RNA transfection and western blot

Protrudin was depleted using SMARTpool: ON-TARGETplus Human ZFYVE27 (118813) small interfering RNA (siRNA) (SMARTpool, L-016349-01-0005, Horizon). The negative siRNA control (MISSION siRNA Universal Negative Control, SIC001) was purchased from Sigma-Aldrich. COS-7 cells were plated in both glass-bottom Petri dishes (for imaging) and six-well plates (for western blot validation). Cells were transfected with 20 nM siRNA oligonucleotides and 20 nM negative control siRNA using Lipofectamine RNAiMax (13778075, Thermo Fisher Scientific) according to the manufacturer's protocol. After 6 h of siRNA transfection, cells were washed, and the medium was replaced with complete culture medium. Twenty-four hours after siRNA transfection, cells were transfected with the plasmid DNA indicated in the Results using Lipofectamine 2000 (Invitrogen). On the day of imaging, cells were stained with SiR-lysosome. Cells in glass Petri dishes were imaged 24 h after DNA transfection.

Cells in six-well plates were collected for western blot validation 72 h after siRNA transfection. Protein concentration was measured using a bicinchoninic acid protein assay kit. Immunoblotting was performed by standard SDS–polyacrylamide gel electrophoresis–western protocols. Primary antibody concentrations were as follows: anti-protrudin at 1:5,000 (Proteintech, 12680-1-AP, R34447), anti-glyceraldehyde-3-phosphate dehydrogenase at 1:30,000 (G8795, Sigma-Aldrich); secondary antibodies (Amersham ECL Rabbit IgG, HRP-linked whole antibody, NA934, 17457635, GE Healthcare Life Sciences; Amersham ECL Mouse IgG, HRP-linked whole antibody (NA931VS, 17234832, GE Healthcare Life Sciences)) were used at 1:3,000 for all rabbit antibodies and for all mouse antibodies. The signal was detected with SuperSignal West Pico Chemiluminescent Substrate.

Widefield and structured illumination microscopy

SIM imaging was performed using a custom three-color system built around an Olympus IX71 microscope stage, which we have previously described[43]. Laser wavelengths of 488 nm (iBEAM-SMART-488, Toptica), 561 nm (OBIS 561, Coherent) and 640 nm (MLD 640, Cobolt) were used to excite fluorescence in samples. The laser beam was expanded to fill the display of a ferroelectric binary spatial light modulator (SXGA-3DM, Forth Dimension Displays) to pattern the light with a grating structure. Polarization of light was controlled with a Pockels cell (M350-80-01, Conoptics). A 60×, 1.2-numerical aperture (NA) water-immersion lens (UPLSAPO 60XW, Olympus) focused the structured illumination pattern onto the sample. This lens also captured the samples' fluorescent emission light before imaging onto an sCMOS camera (C11440, Hamamatsu). The maximum laser intensity at the sample was 20 W cm^{-2}. Widefield images and raw SIM images were acquired with HCImage Live software (Hamamatsu) to record image data to a disk and a custom LabVIEW 2016 program (freely available upon request) to synchronize the acquisition hardware. Multicolor images were registered by characterizing channel displacement using a matrix generated with TetraSpeck beads (Life Technologies) imaged in the same experiment as the cells. COS-7 cells expressing mEmerald-Sec61b-C1 (ER marker) and stained with SiR-lysosome (lysosome marker) were imaged by SIM every 1.5 s (including imaging exposure time (20–30 ms for each channel) of both channels) for 60 frames.

Reconstruction of SIM images with LAG SIM

Resolution-enhanced images were reconstructed from raw SIM data with LAG SIM, a custom plugin for Fiji–ImageJ available in the Fiji Updater. LAG SIM provides an interface to the Java functions provided by fairSIM[44]. LAG SIM allows users of our custom microscope to quickly iterate through various algorithm input parameters to reproduce SIM images with minimal artifacts; integration with SQUIRREL[45] provides numerical assessment of such reconstruction artifacts. Furthermore, once appropriate reconstruction parameters have been calculated, LAG SIM provides batch reconstruction of data so that a folder of multicolor, multi-frame SIM data can be reconstructed overnight with no user input.

Airyscan imaging

Airyscan imaging was performed using an LSM 880 confocal microscope (Zeiss). A Zeiss Plan-Apochromat 63×, 1.40-NA DIC M27 oil objective was used. For visualization of ER structure, ER mCherry was excited by a diode-pumped solid-state 561-nm laser (1% intensity) and detected using the Airyscan detector. Bit depth was set at 16 bits. Using the Fast-Airyscan mode, live cell time-lapse images were acquired every 1 s (60 frames) with an image size of 1,364 × 1,244 pixels. Cells were kept in a controlled environment (37 °C, 5% CO_2) during imaging. Following acquisition, images were deconvoluted using Airyscan processing. Image processing was performed with ZEN 2.3 SP1 FP3 (black) software (version 14.0.25.201).

Confocal imaging

A part of confocal imaging was performed using a STELLARIS 8 confocal microscope (Leica). An HC PL APO CS2 63×, 1.40-NA oil objective was used. For visualization of ER structure, ER mCherry was excited with a white light laser at 587 nm with 3% intensity and detected using the HyD S3 detector (detection range, 592–750 nm). Bit depth was set at 16 bits. Live cell time-lapse images were acquired every 1.5 s (90 frames) with an image size of 512 × 512 pixels. Cells were kept in a controlled environment (37 °C, 5% CO_2) during imaging.

ERnet construction and training

For segmentation of sequential ER images, a spatio-temporal shifted window vision transformer neural network was trained and used. The proposed model was inspired by the previous models Vision Transformer[19], its more efficient shifted window variant Swin[20] and adaptation to image restoration with SwinIR[46]. We also combined the MSA mechanism[33] with a channel attention mechanism[34] in ERnet, a design that makes the model more adaptive to different phenotypes of the ER. Swin introduced the inductive bias to self-attention called SW-MSA, which can be compared to the inductive bias inherent in convolutional networks. SwinIR introduced residual blocks to the Swin Transformer to help preserve high-frequency information for deep feature extraction. The Video Swin Transformer extended the SW-MSA to three dimensions, such that spatio-temporal data could be included in the local attention for the self-attention calculation. Further to this, the success of the channel attention mechanism[47] inspired the inclusion of this other inductive bias in addition to 3D local self-attention following the SW-MSA approach.

The inputs to the model have the dimension $T \times H \times W \times C$, where H is the height, W is the width of the image dimensions, T is 5 for ERnet (five adjacent temporal frames) and C is 1 (grayscale inputs). A shallow feature-extraction module in the beginning of the network architecture, shown in Fig. 1, projects the input into a feature map, F_0, of the $T \times H \times W \times D$ dimension, where the embedding dimension D is a hyperparameter. The feature map is passed through a sequence of residual blocks denoted as WCAB:

$$F_i = H_{\mathrm{WCAB}}(F_{i-1}), \quad i = 11, .., n\,.$$

Inside each WCAB is a sequence of STLs, in which MSA is calculated using local attention with a shifted window mechanism. Inputs to the STL layer are partitioned into $\frac{T}{P} \times \frac{HW}{M^2}$ 3D tokens of the $P \times M^2 \times D$ dimension, where P and M are the patch sizes along the temporal and spatial axes, respectively. Here, we use $T = 3$ and $M = 4$. For a local window feature, $x \in \mathbb{R}^{P \times M^2 \times D}$, query ($Q$), key ($K$) and value ($V$) matrices, $\{Q, K, V\} \in \mathbb{R}^{PM^2 \times D}$, are computed by multiplication with projection matrices following the original formulation of transformers. Attention is then computed as

$$\mathrm{Attention}(Q, K, V) = \mathrm{SoftMax}(QK^T/\sqrt{d} + B)V,$$

where d is the dimension of query and key features and $B \in \mathbb{R}^{P^2 \times M^2 \times M^2}$ is a relative positional bias found to lead to substantial improvements in classification performance. STLs are joined in a way similar to the residual blocks, although the use of SW-MSA is alternated with a version without shifted windows, W-MSA, ensuring that attention is computed across window boundaries, which would not have been the case without SW-MSA.

After the final STL, the mth layer in a WCAB, a transposed 3D convolutional layer is used to project the 3D tokens back into a $T \times H \times W \times D$ feature map, $F_{i,m}$. A channel attention module is then used on $F_{i,m}$ to determine the dependencies between channels following the calculation of the channel attention statistic. The mechanism works by using global adaptive average pooling to reduce the feature map to a vector, which, after passing through a 2D convolutional layer, becomes weights that are multiplied back onto $F_{i,m}$ such that channels are adaptively weighed. A residual is then obtained by adding a skip connection from the beginning of the ith WCAB to prevent loss of information, that is, low-frequency information, and the vanishing gradient problem. A fusion layer combines the temporal dimension and the channel dimensions. For the final upsampling module, we use the subpixel convolutional filter to expand the image dimensions by aggregating the fused feature maps.

The model is trained by minimizing a multi-class cross-entropy loss function,

$$L_{\mathrm{CE}}(\Theta; D) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{WH}\sum_{x=1}^{W}\sum_{y=1}^{H}\sum_{k=1}^{K} -f^{H}_{i;x,y}(k)\log\left[\frac{\exp\left(F\left(\Theta; I^{L}_{i}\right)_{x,y;k}\right)}{\sum_{j=1}^{K}\exp\left(F\left(\Theta; I^{L}_{i}\right)_{x,y;j}\right)}\right]\right),$$

where k and j are iterators over a total of K unique classes, and $f^{H}_{i;x,y}(k)$ is a function equal to 1 if the target class for the pixel at (x, y) of the ith image is k and equal to 0 otherwise. In this paper, we study the segmentation of background, tubules, sheets and SBTs, and, therefore, $K = 4$ in the equation above. Θ represents the trainable parameters of the network referred to as $\mathrm{F}(\cdot)$, while D is the training dataset of size N consisting of the set of image samples I^{L}_{i}. I refers to input, L refers to low quality images.

The training data are obtained by acquiring experimental data using SIM. A total of 20 sequential stacks of different samples are acquired, where each stack consists of 60 SIM images reconstructed with ML-SIM. The super-resolved SIM outputs are then segmented by manually fine-tuning a random forest model in the Weka plugin for ImageJ on an image-by-image basis.

ERnet has been trained with the Adam optimizer and a cross-entropy loss function using a learning rate of 1×10^{-5} that is halved after 30,000 iterations. A total of 65,000 iterations were made, which equals 100 epochs of the training dataset. A Nvidia A100 GPU was used with a batch size of ten. Training samples were randomly cropped to 128×128, while inference was performed with $1{,}024 \times 1{,}024$ inputs. For ERnet, the WCAB number, STL number, window size, embedding size D and attention head number are set to 6, 6, 8, 96 and 6, respectively. The other hyperparameters are further specified below.

Implementation details are as follows: implementation, Training/archs/swin3d_rcab3_arch.py; patch size, (3, 4, 4); window size, (2, 8, 8); MLP ratio, 2; number of STLs, 5; depths of STLs, (6, 6, 6, 6, 6); embedding dimension, 192; attention head number, (8, 8, 8, 8, 8); batch size, 10; image size, 128; number of input channels, 1; number of output channels, 4; data workers, 4; validation images, 70; training images, 650; number of epochs, 100; learning rate, 0.0001; learning scheduler, reduced by 0.5 per 50 epochs.

Network analysis methods

To quantify structural changes in the ER, methods from network analysis are applied[48,49]. We represent the ER structure of tubules through an undirected and unweighted graph. All tubule junctions are represented by nodes, and tubules are represented by edges.

Networks are built in a Python routine, and their metrics are measured through the Python package graph-tool[18] and NetworkX[50]. We measure the size of the network through the number of nodes N and edges E within the system. The number of edges attached to one node is called the node degree k, and the distribution of the degrees is one of the most fundamental parts of the analysis of network structures.

To quantify structural arrangements of the ER, we focus on primary network connectivity metrics. First, we measure the network density d between nodes and edges (equation (2)). Other metrics that describe network connectivity are the global clustering coefficient (equation (2)) and network assortativity (equation (3)). The global clustering coefficient describes the tendency of the network to build triangles by relating triplets to each other. Three nodes connected to each other through three edges are a 'closed triplet', while three nodes connected to each other through two edges are called an 'open triplet' (ref. 51). Network assortativity describes the likelihood of nodes connecting with nodes of similar properties, here specifically, as is common, a node degree. Assortative mixing is contrasted with disassortative mixing where nodes tend to connect to others of dissimilar properties[52]. The assortativity coefficient r is described in equation (3), where e_{ij} is the fraction of edges linking a node with type i to nodes of type j, a_i is the sum over e_{ij} for all j, and b_i is the sum over e_{ij} for all i. An assortativity coefficient of $r = 0$ indicates no mixing preference,

whereas positive values indicate assortative tendencies and negative values indicate disassortative tendencies.

$$d = \frac{2E}{N(N-1)} \quad (1)$$

$$Cl = \frac{\text{number of closed triplets}}{\text{number of all triplets}} \quad (2)$$

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i} \quad (3)$$

Additionally, we include macroscopic network arrangements by counting the number of network components. Networks may be entirely connected or composed of many distinct components[53]. For networks evolving over time, network components outline merging or splitting behavior. In networks with many components, the most characteristic topological features are often exhibited in the largest component[54].

Ground truth test of connectivity analysis

First, we generate a random network and use triangulation and tessellation to obtain a fully connected network. Using cubic spine interpolation (third panel), we generate a backbone that mimics a connected ER tubular network. This dataset can then be processed to mimic microscopic imaging data through addition of noise and PSF blurring. The noise level is a parameter defined here as a scaling factor of the standard deviation of a Gaussian noise source, ranging from 0 to 20. The SNR values follow a more standardized definition given by the ratio of the mean of the signal and the standard deviation of the background. The SNR for random noise N is defined as

$$\text{SNR} = \frac{\boldsymbol{E}[\boldsymbol{S}^2]}{\boldsymbol{E}[\boldsymbol{N}^2]}.$$

S is a random variable representing the signal. S^2 is the signal power. If the noise has the expected value of zero, the denominator is its variance, the square of its standard deviation σ_N.

Data visualization

Videos of time-lapse imaging and analysis were performed using Fiji (NIH). Connectivity graphs in the figures were replotted by a Python module named 'connectivity graph.py'. Instructions for using this module are provided inside the file. Colors of the segmented ER domains, including tubules, sheets and SBTs, are displayed in grayscale format from ERnet, which can be changed based on the user's preference.

Statistical analysis

Statistical significance between two values was determined using a two-tailed, unpaired Student's t-test (GraphPad Prism 8.2.1). Statistical analysis of three or more values was performed by one-way ANOVA with Tukey's post hoc test (GraphPad Prism). All data are presented as mean ± s.e.m.; $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$.

Statistical parameters including the exact value of n, the mean, median, dispersion and precision measures (mean ± s.e.m.) and statistical significance are reported in the figures and figure legends. Data were judged to be statistically significant when $P < 0.05$ by two-tailed Student's t-test. In the figures, asterisks denote statistical significance as calculated by Student's t-test ($^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$).

Reporting summary

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

Data availability

All data needed to evaluate the conclusions in the paper are present in the Source data. All the datasets used to train and test the model are publicly accessible at the figshare repository: https://figshare.com/articles/dataset/ERnet_datasets/21975878/1. Source data are provided with this paper.

Code availability

The ERnet model is written in Python. The software and Colab versions of ERnet are also freely available online through GitHub at https://github.com/charlesnchr/ERnet-v2.

References

39. Michel, C. H. et al. Extracellular monomeric tau protein is sufficient to initiate the spread of tau protein pathology. *J. Biol. Chem.* **289**, 956–967 (2014).
40. Avezov, E. et al. Retarded PDI diffusion and a reductive shift in poise of the calcium depleted endoplasmic reticulum. *BMC Biol.* **13**, 2 (2015).
41. Middya, S. et al. Microelectrode arrays for simultaneous electrophysiology and advanced optical microscopy. *Adv. Sci.* **8**, 2004434 (2021).
42. Lu, M. et al. Structural progression of amyloid-β Arctic mutant aggregation in cells revealed by multiparametric imaging. *J. Biol. Chem.* **294**, 1478–1487 (2019).
43. Young, L. J., Ströhl, F. & Kaminski, C. F. A guide to structured illumination TIRF microscopy at high speed with multiple colors. *J. Vis. Exp.* https://doi.org/10.3791/53988 e53988 (2016).
44. Müller, M., Mönkemöller, V., Hennig, S., Hübner, W. & Huser, T. Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ. *Nat. Commun.* **7**, 10980 (2016).
45. Culley, S. et al. Quantitative mapping and minimization of super-resolution optical imaging artifacts. *Nat. Methods* **15**, 263–266 (2018).
46. Liang, J. et al. SwinIR: image restoration using Swin Transformer. In *Proc. IEEE/CVF International Conference on Computer Vision* 1833–1844 (ICCV, 2021).
47. Zhang, Y. et al. Image super-resolution using very deep residual channel attention networks. In *Proc. European Conference on Computer Vision* 286–301 (ECCV, 2018).
48. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D. U. Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308 (2006).
49. Costa, L. et al. Characterization of complex networks: a survey of measurements. *Am. J. Enol. Vitic.* **38**, 293–297 (1987).
50. Hagberg, A., Swart, P. & Chult, S. D. *Exploring Network Structure, Dynamics, and Function Using NetworkX*. Report No. LA-UR-08-05495 (USDOE, 2008).
51. Newman, M. E. J. The structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003).
52. Cimini, G. et al. The statistical physics of real-world networks. *Nat. Rev. Phys.* **1**, 58–71 (2019).
53. Albert, R. Scale-free networks in cell biology. *J. Cell Sci.* **118**, 4947–4957 (2005).
54. Strogatz, S. H. Exploring complex networks. *Nature* **410**, 268–276 (2001).

Acknowledgements

We thank A.I. Fernández Villegas and Y. Feng for helping with cell culture. We thank E. Ward for helping with image processing. We thank J. Hu (Chinese Academy of Sciences, China) for giving us the ATL KO cell line. Funding: this research was funded by Infinitus (China) (supporting M.L., C.F.K. and G.S.K.S.); a Wellcome Trust Programme Grant (085314/Z/08/Z to G.S.K.S. and C.F.K);

a Swiss National Science Foundation Career Grant (P2EZP2_199843 to N.F.L.); a research fellowship from the Deutsche Forschungsgemeinschaft (DFG; SCHE 1672/2-1 to K.M.S.) and pump-prime funding from the Integrated Biological Imaging Network (G106925 to K.M.S.); the UK Dementia Research Institute, which receives its funding from the UK DRI, funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK (supporting T.K., E.A. and C.F.K.) and Alzheimer's Society 525 (AS-PhD-19a-015) supporting E.A. J.M.W.'s PhD scholarship was funded by the Department of Chemical Engineering and Biotechnology, University of Cambridge.

Author contributions

M.L. designed, conducted and interpreted experiments and wrote the article. M.L. and C.N.C. developed the computational pipeline for ERnet. C.N.C. developed the core model of ERnet. J.M.W. conceptualized, developed and wrote the graph-based analysis of the ER. T.K. supported the versatility test. N.F.L., K.M.S., E.A., P.L., A.A.L. and G.S.K.S. gave advice and edited the article. C.F.K. supervised the research, coordinated and conceptualized the study and wrote the article.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41592-023-01815-0.

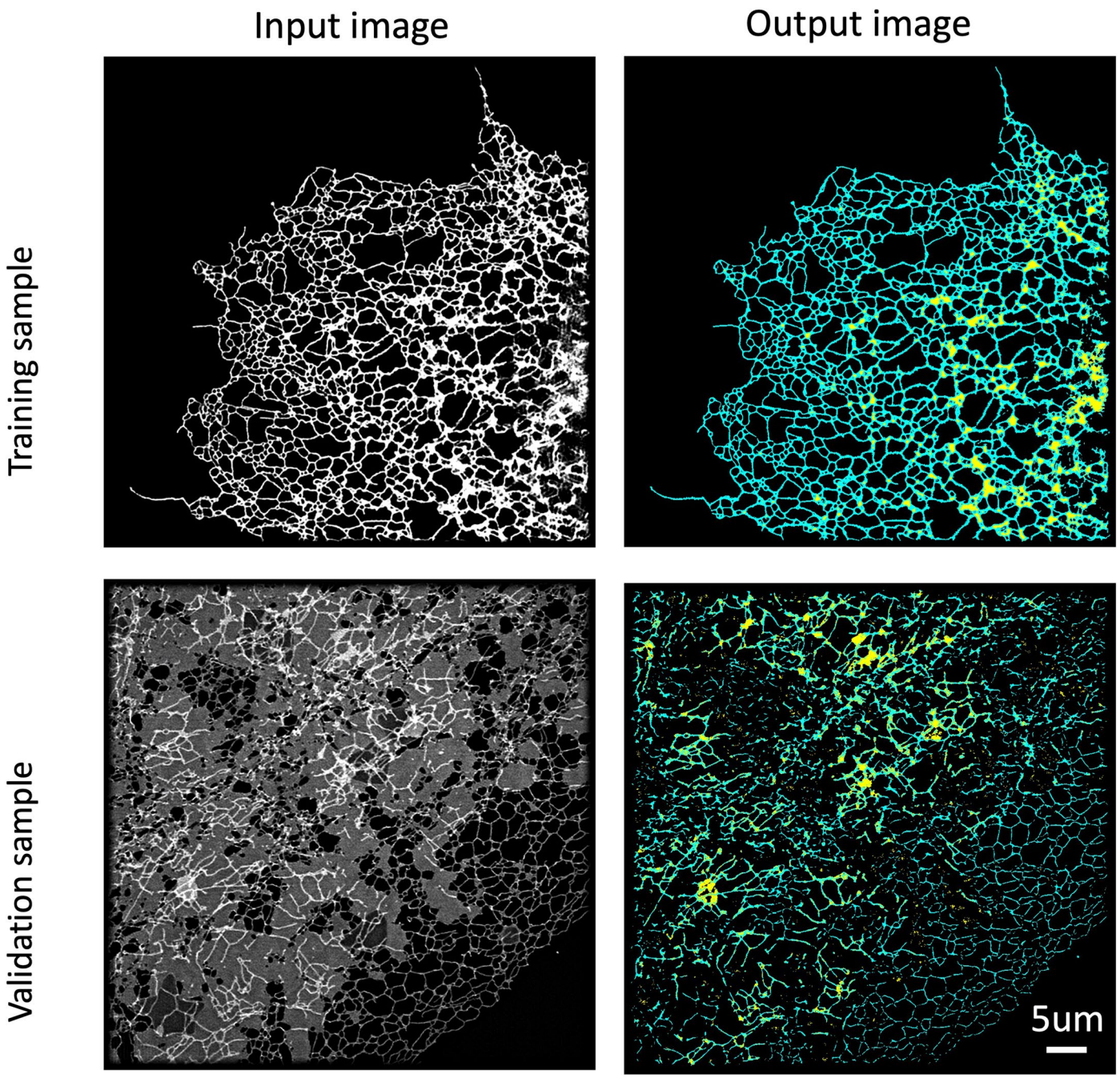
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-023-01815-0.

Correspondence and requests for materials should be addressed to Clemens F. Kaminski.

Peer review information *Nature Methods* thanks Niamh O'Sullivan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Rita Strack, in collaboration with the *Nature Methods* team. Peer reviewer reports are available.

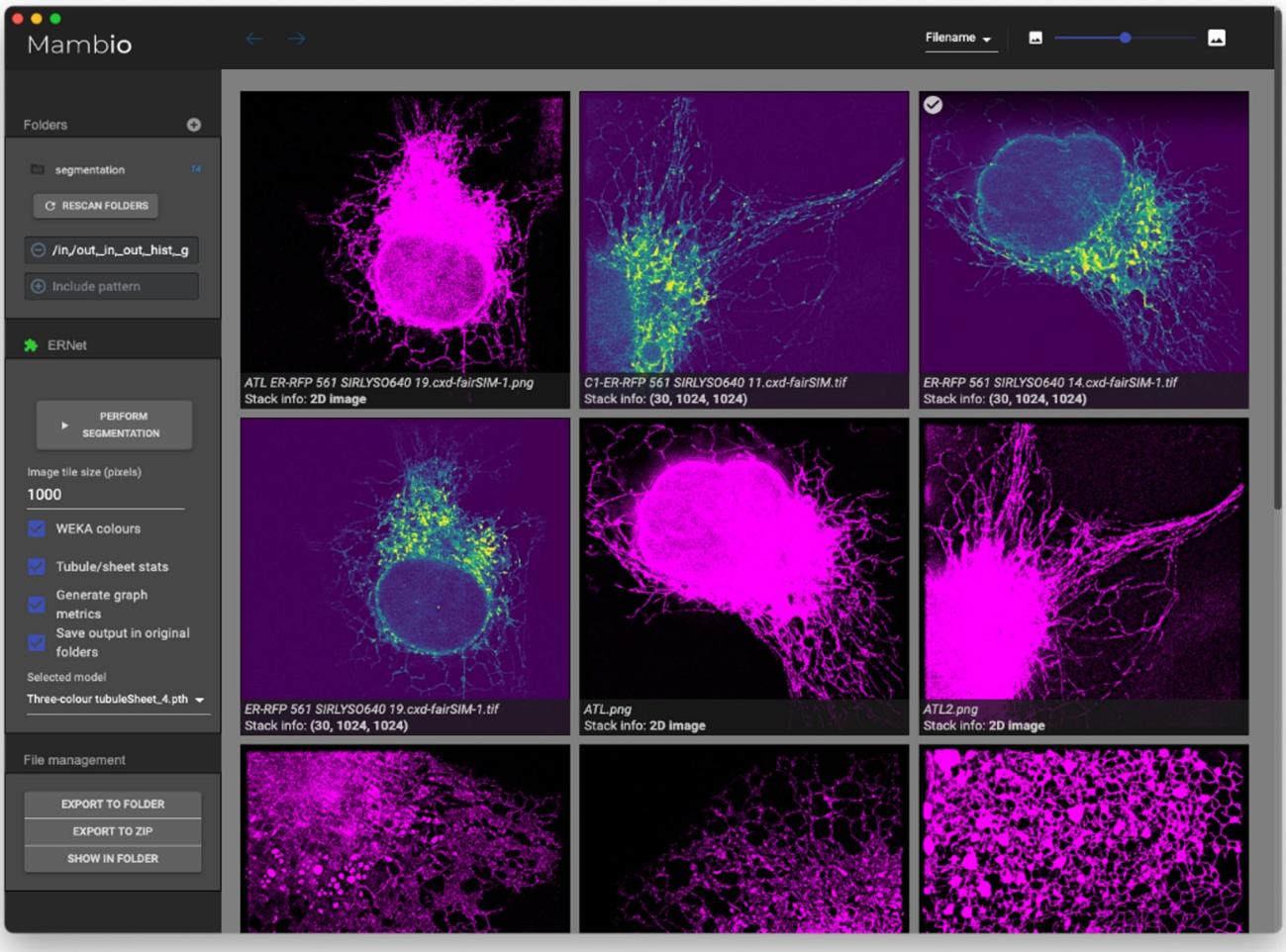
Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41592-023-01815-0



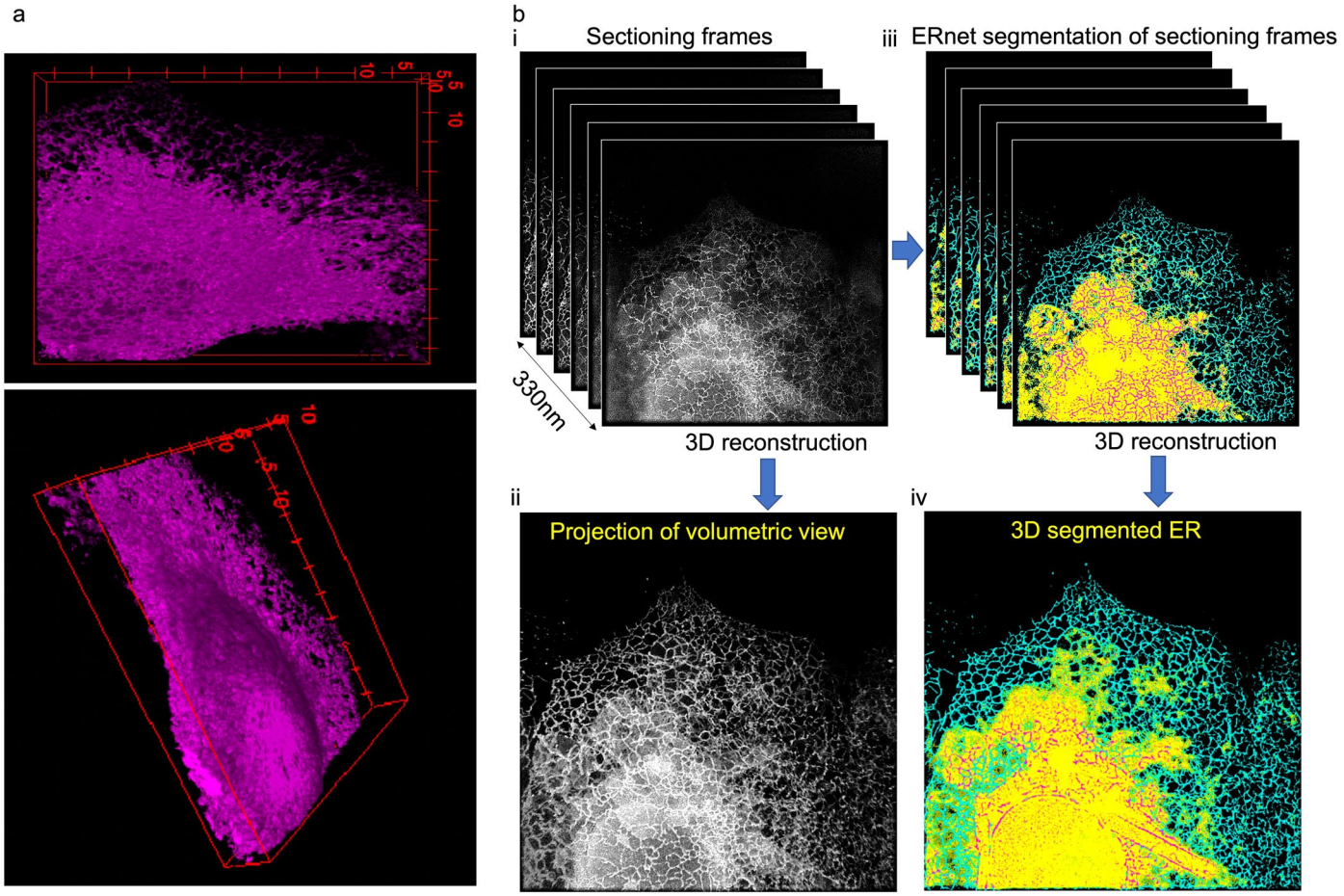
Extended Data Fig. 1 | A test of Weka trainable segmentation with different input data. Top left: An input image was used to train a classifier of Weka Trainable Segmentation. Top right: The tubules (cyan) and sheet (yellow) can be clearly classified after segmentation. Bottom left: a new image was applied to the trained classifier shown above. Bottom right: segmentation result of the new input data. This test was independently repeated five times with similar results. Scale bars: 5 μm.

https://doi.org/10.1038/s41592-023-01815-0



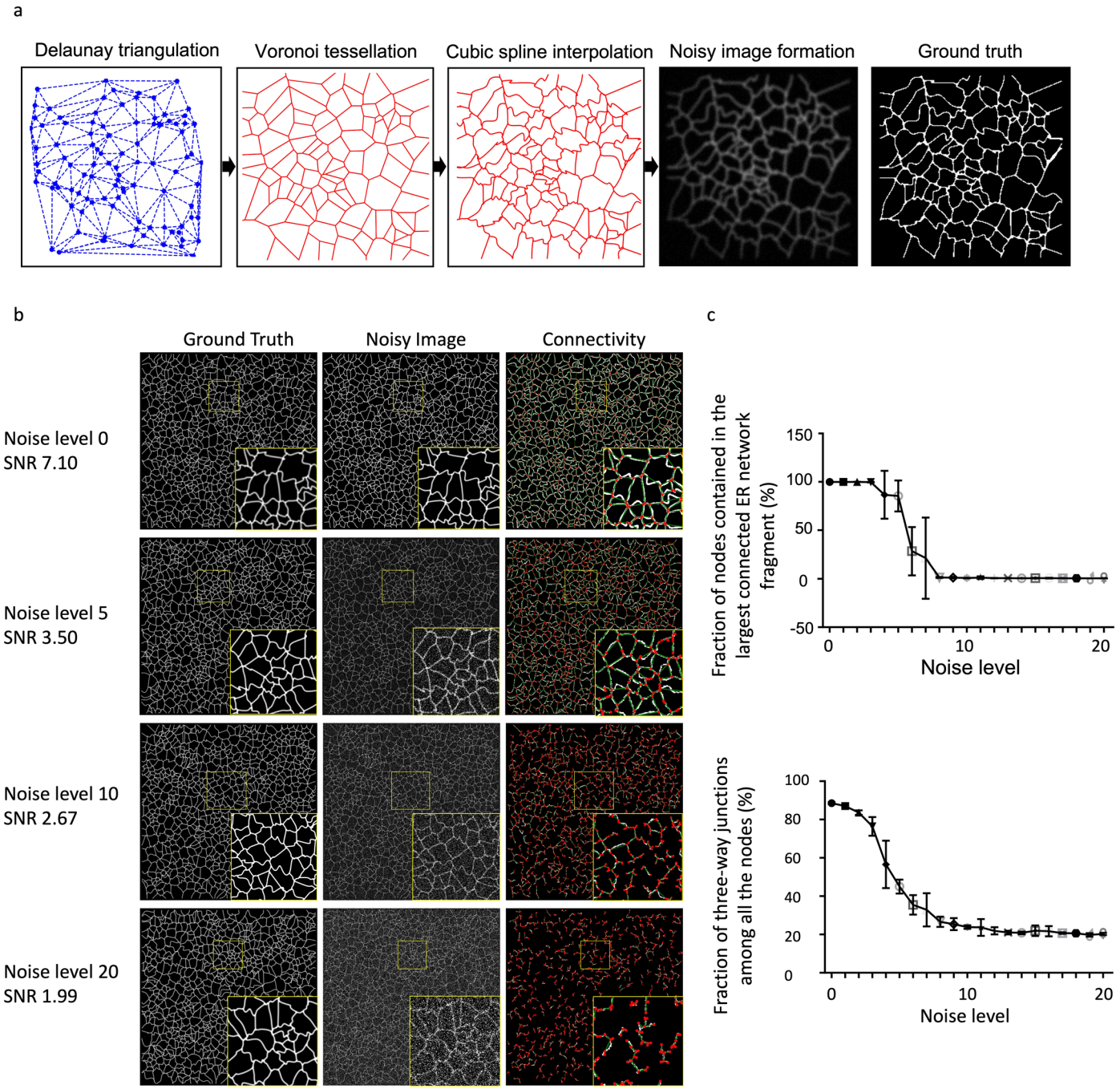
Extended Data Fig. 2 | ERnet graphical user interface. Left part of the interface shows the path of input and output images. Bottom left: options of the analysis provided by ERnet. Right part of the interface shows the input images (magenta) and segmented results.

https://doi.org/10.1038/s41592-023-01815-0



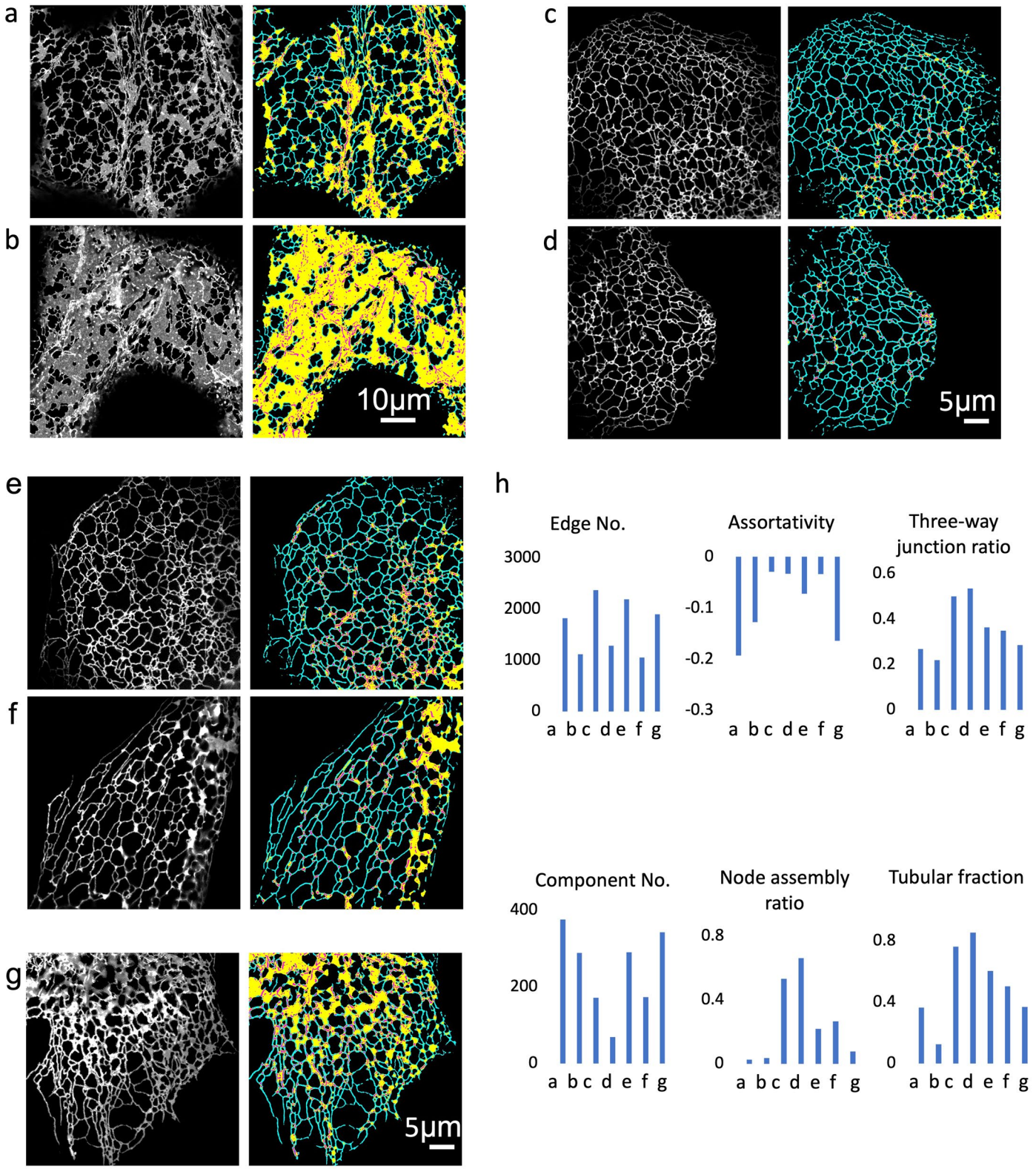
Extended Data Fig. 3 | Validation of 2D analysis of 3D ER structure. (a) 3D projection view of a U2OS cell expressing sec61-mCherry. The peripherical ER is mainly distributed as a flat, single layer while the perinuclear ER is more 3 dimensional in character (see Supplementary Video 12). (**b**) 3D projection view of a COS-7 cell expressing mEmerald-Sec61b-C1. i: (left) 3D reconstruction of image sections. ii: projection of the volumetric view of corresponding the sections shown in (i). Peripherical ER features are flat and singly layered (Supplementary Video 13). In the perinuclear region, multiple layers of ER are visible. iii: each section from the 3D image stack shown in (i) is segmented individually by ERnet. iv: The resulting segmented frames are then combined into a 3D volume to show the network topology (Supplementary Video 14). Cyan regions (tubules) show flat, 2D regions with little 3D character. Perinuclear regions appear in yellow (sheets), where the topology is 3 dimensional in character.

https://doi.org/10.1038/s41592-023-01815-0



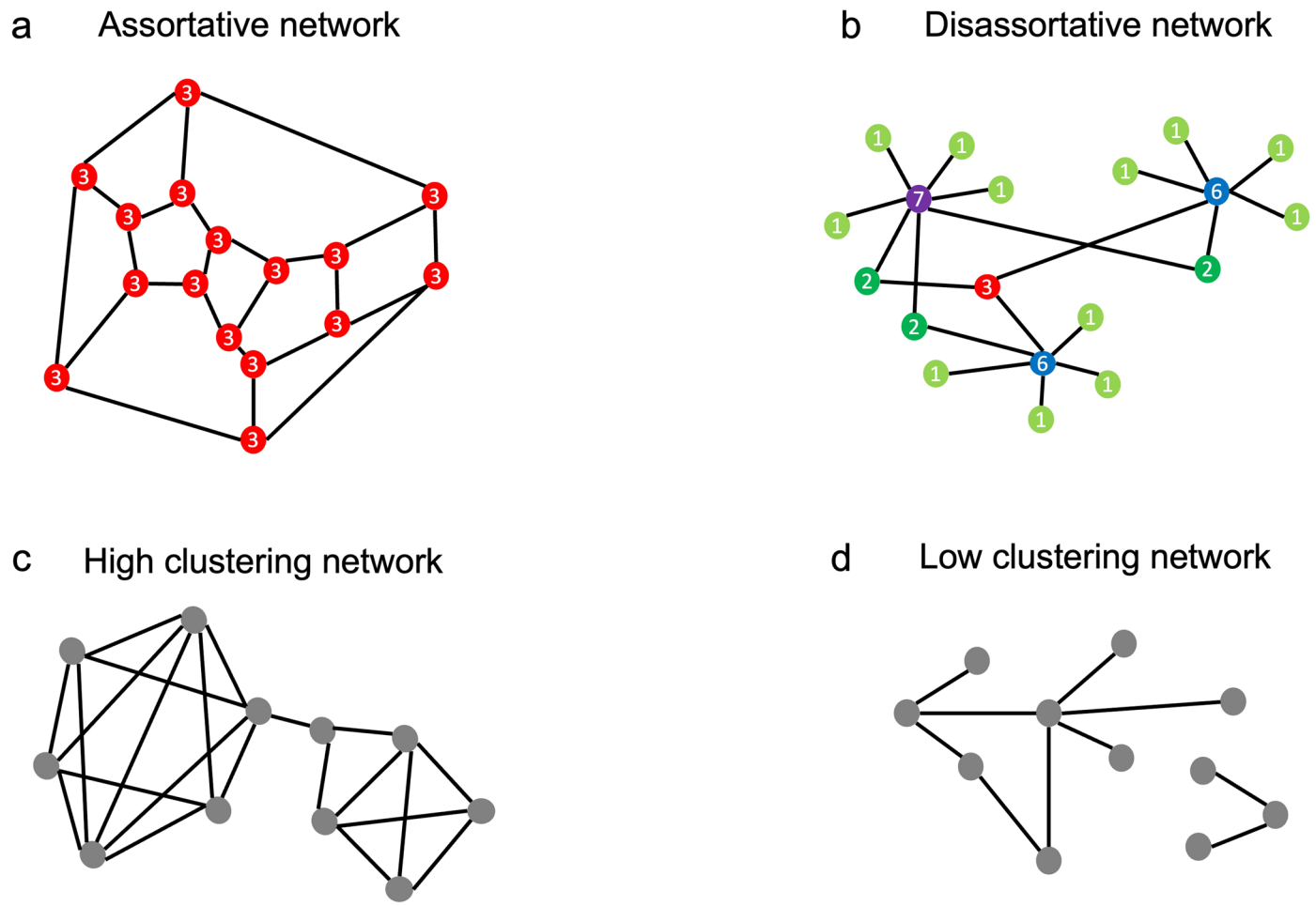
Extended Data Fig. 4 | Performance validation of ERnet as a function of signal-to noise-ratio prevailing in ER image data. (**a**) Pipeline to develop synthetic ground truth data to mimic ER structure as it would appear when imaged under the microscope including resolution loss through PSF blurring, and addition of image noise affecting the Signal to Noise ratio (SNR). A network was generated from points distributed randomly over a given field of view. Triangulation and tessellation were then used before cubic spine interpolation to obtain a network mimicking features of a fully connected tubular ER system. We added Gaussian noise and blurred images with a PSF kernel to produce ground truth data as obtained by our microscope setup. (**b**) Representative synthetic ER networks were generated at different noise levels. Ground truth images (left column) were processed, and Gaussian noise superimposed (middle column). Images were then analysed by ERnet to produce the skeleton map displaying the connectivity (right column), including nodes and edges. The output could then be directly compared with the results from ground truth data. (**c**) Degradation of network metrics as a function of decreasing image signal to noise ratio. The top panel compares the number of nodes obtained from the noisy data with that of the ground truth data (1 = fully matched, 0 = no match) as a function of noise level. The bottom panel shows how the number of three-way junctions (node of degree three) identified by ERnet decreases as noise increases. In the ground truth image (Noise level 0) ca 90% of all nodes constitute three-way junctions. For image signal to noise-ratios exceeding 5, ERnet reproduced the network topology to within 90% of the ground truth. $N = 5$ synthetic images per noise level. Data are presented as mean ± SEM. The algorithms and full instructions to generate synthetic ER networks mimicking microscopy data are available on GitHub and see Source Data Extended Data Fig. 4c. (https://github.com/charlesnchr/ERnet-v2).

https://doi.org/10.1038/s41592-023-01815-0



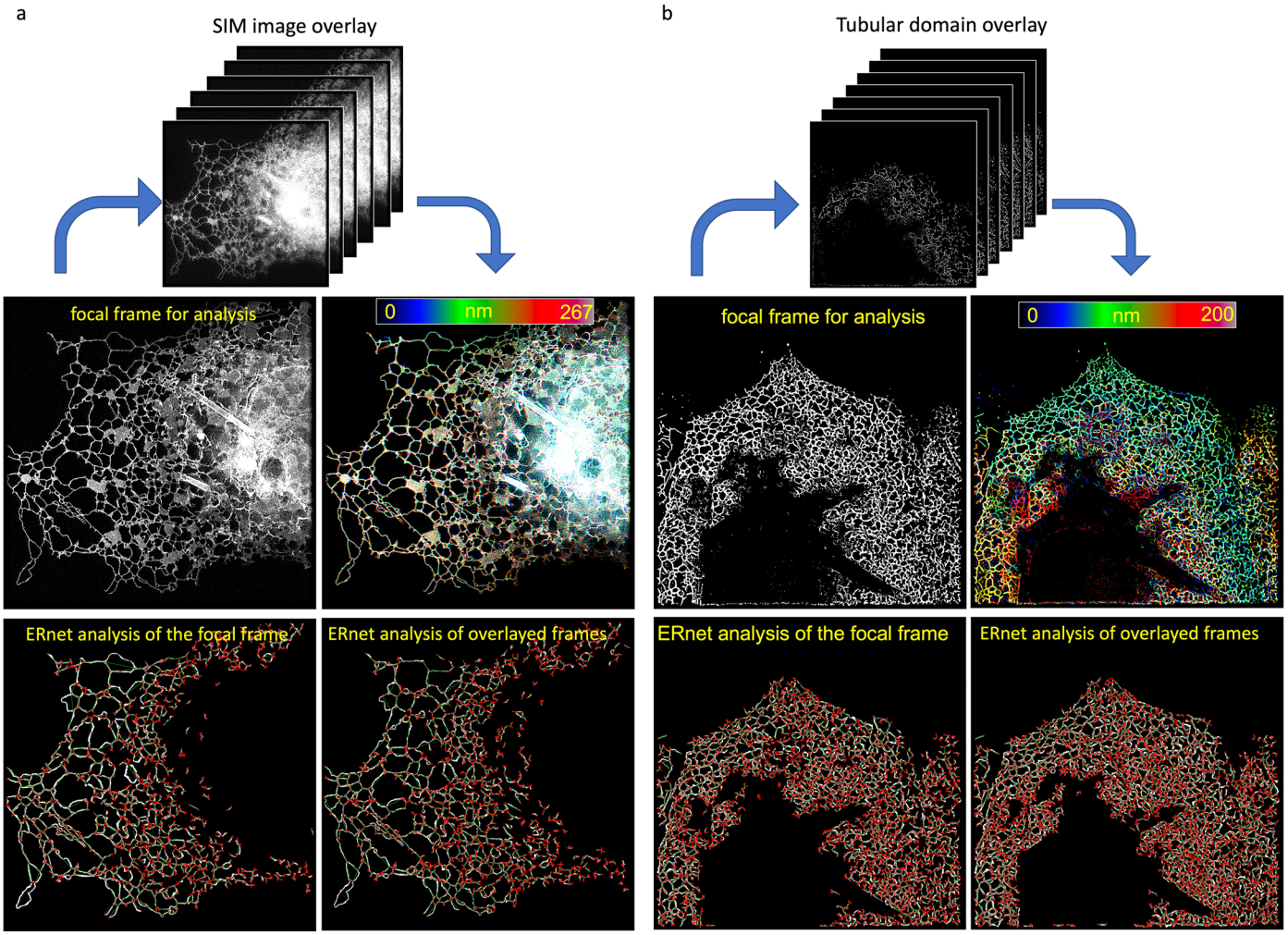
Extended Data Fig. 5 | Validation of ERnet on publicly available datasets of ER images. ERnet recovers network topologies from ER recorded with different imaging techniques and cell types, including mammalian and plant cells. (**a**)–(**g**) Raw images (grayscale images, left hand side) and corresponding results from ERnet image segmentation (colour images, right hand side). Tubules (cyan), sheets (yellow), and sheet-based tubules, SBTs (magenta) are clearly recovered in all cases studied. (a-b) are from the dataset provided by Pain et al.[13] and represent ER present in plant cells recorded with AiryScan confocal imaging; 33 different images from this dataset were reanalysed by ERnet and compared with the results from AnalyzER; (c-f) are from the dataset provided by Qiao et al., 2021 (COS-7 cells, recorded with SIM); 18 different images from this dataset were tested by ERnet; (g) is from the dataset provided by Qin et al., 2020 (COS-7 cells, recorded with SIM); 2 different images from this dataset were tested by ERnet; (**h**) ER morphology metrics for all structures identified by ERnet. See Source Data Extended Data Fig. 5.

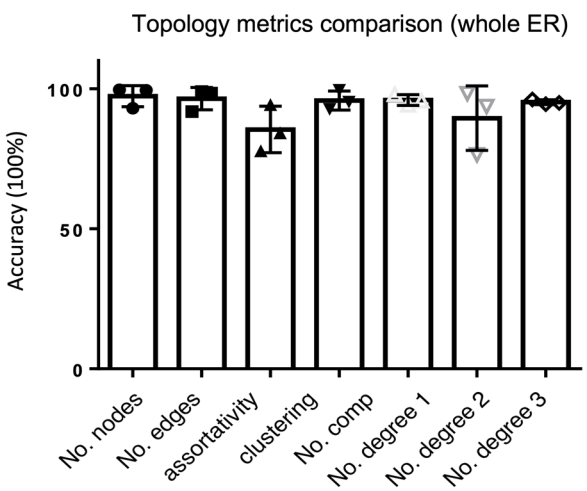
https://doi.org/10.1038/s41592-023-01815-0

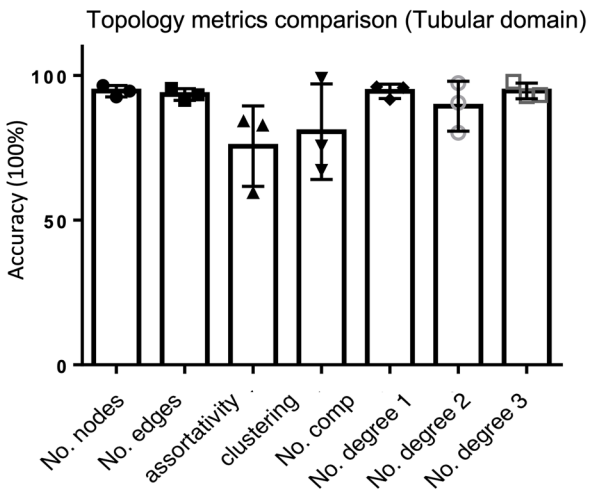


Extended Data Fig. 6 | Graphical explanation of graph theory concepts. (**a**) In this example of an assortative network, nodes show a preference of attaching to others that are similar. Here, nodes with the same degree tend to attach to each other (a node's degree is represented by its colour and number in the node). (**b**) In this example of disassortative network, nodes show a preference of attaching to others that are not similar. Here, nodes with different degrees tend to attach to each other. (**c**) In this example, a network is shown characterised by a high clustering coefficient: here nodes tend to connect also to the neighbours of their neighbours, leading to the formation of network triangles. (**d**) Example of a network with a low clustering coefficient: nodes do not tend to connect to the neighbours of their neighbours. None, or only few, network triangles are formed as a result.

https://doi.org/10.1038/s41592-023-01815-0

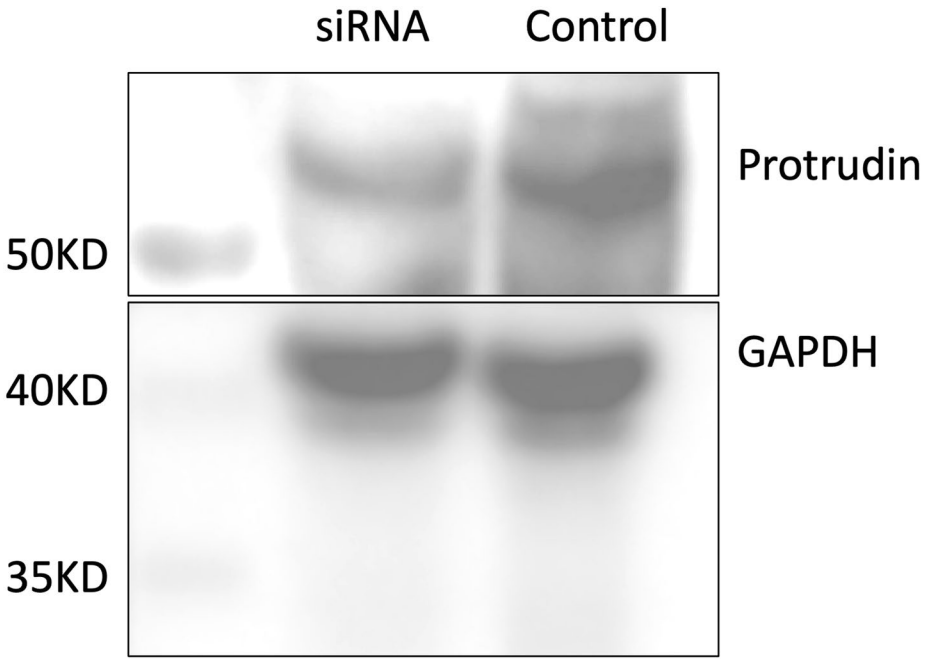






Extended Data Fig. 7 | Quantitative comparison between the 2D and 3D analysis. (**a**) Top panel: 3D SIM image sections were overlayed to produce a flattened image of the ER; left: central section of the ER image stack; right: colour coded flattened ER structures. Bottom: comparison of ERnet results for both types of images reveal nearly identical connectivity, supporting the notion that a 2D analysis suffices to capture ER topology in the cell types studied. Red spots: nodes; green lines: edges. Lower bottom: comparison of the connectivity metrics. To do this, we obtained volumetric sections of ER structures with SIM and reconstructed the data in 3D. We then applied ERnet on flattened image stacks or individual 2D sections and compared the two results. For phenotyping the effect of drugs in cell lines such as COS-7 or U2OS, which are morphologically flat, the advantages in speed and applicability of a 2D method greatly outweigh the disadvantage of not capturing the occasionally occurring 3-dimensional network features. $N = 3$ sectioning SIM images. Data are presented as mean ± SEM. See Source Data Extended Data Fig. 7a. (**b**) Same analysis to (a) performed on tubular domains only. $N = 3$ sectioning SIM images. Data are presented as mean ± SEM. See Source Data Extended Data Fig. 7b.

https://doi.org/10.1038/s41592-023-01815-0



Extended Data Fig. 8 | Western blot validation of Protrudin depletion. Validation of siRNA depletion of Protrudin by Western blot. Two independent experiments were repeated with similar results. See source data file for uncropped scans of blots.

nature portfolio

Corresponding author(s): Clemens Kaminski

Last updated by author(s): Feb 7, 2023

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	Confocal images were acquired through STELLARIS 8 confocal microscope software (Leica). AiryScan images were acquired through software ZEN 2.3 SP1 FP3 (ver.14.0.25.201). Widefield images and raw SIM images were acquired with the HCImage Live software (Hamamatsu) to record image data to disk and a custom LabView 2016 program (freely available upon request) to synchronize the acquisition hardware.
Data analysis	AiryScan image processing was performed in software ZEN 2.3 SP1 FP3 (ver.14.0.25.201). SIM images were reconstructed from the raw SIM data with LAG SIM, a custom plugin for Fiji/ImageJ available in the Fiji Updater. Images were processed through Fiji (NIH). ER structures were analysed through ERnet https://github.com/charlesnchr/ERnet-v2. The connectivity graphs in the figures are re-plotted through a Python module named "connectivity graph.py", https://github.com/charlesnchr/ERnet-v2. Statistical significance was analysed through Prism-GraphPad 8.2.1.

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

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data needed to evaluate the conclusions in the paper are present in the Source Data files. All the datasets used to train and test the model are publicly accessible at figshare repository: https://figshare.com/articles/dataset/ERnet_datasets/21975878/1.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	500 cells were examined to test the presence of sheet-based tubules. This sample size was as large as could be practically attained. Each analysis of ER structure through ERnet was conducted based on single cell sequential images with at least 30 frames recorded. More than 10 cells, a number exceeding the typical numbers in the field, were analysed through ERnet for each condition described in figure legends in manuscript. The sample sizes were chosen based on our previous study (Lu et al., 2020).
Data exclusions	No data were excluded from analyses presented in this study.
Replication	All attempts at replication were successful. All experiments were repeated at least three times. A large versatility test on different cell lines and microscopy techniques was conducted to validate the performance of ERnet and the result replication.
Randomization	No randomisation was performed. Samples were allocated into groups based on specific treatments. Randomisation in these experiments were not feasible.
Blinding	The investigators were blinded to group allocation during data collection. The analysis for experiments comparing different conditions is blinded as the ERnet model takes no subjective user input.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used

Protrudin antibody (Proteintech, Cat#12680-1-AP, Lot R34447), GAPDH antibody (Sigma-Aldrich, Cat#G8795), Amersham ECL Rabbit IgG, HRP-linked whole antibody (NA934, Lot 17457635, GE Healthcare Life Sciences), Amersham ECL Mouse IgG, HRP-linked whole antibody (NA931VS, Lot 17234832, GE Healthcare Life Sciences).

Validation

Protrudin antibody was validated via western blot analysis by Proteintech in mouse brain tissue, Jurkat cells, mouse testis tissue, and mouse thymus tissue.
GAPDH antibody was validated via western blot analysis by Sigma-Aldrich in NIH-3T3, human Hela, Jurkat, COS-7, PC-12, Rat2, CHO, MDBK and MDCK cells.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

All the cell lines were purchased from American Culture Collection (ATCC): COS-7 cells (CRL-1651, ATCC), SH-SY5Y cells (CRL-2266, ATCC), HEK293T cells (CRL-3216, ATCC), CHO-K1 (CCL-61, ATCC), U2OS cells (HTB-96, ATCC).

Authentication

Cell lines were not further authenticated.

Mycoplasma contamination

The cell lines were confirmed to be mycoplasma free.

Commonly misidentified lines
(See ICLAC register)

None.