



Automatic segmentation and reconstruction of intracellular compartments in volumetric electron microscopy data

Manca Žerovnik Mekuč^a, Ciril Bohak^{a,c}, Eva Boneš^a, Samo Hudoklin^b, Rok Romih^b, Matija Marolt^{a,a,*}

^a Faculty of Computer and Information Science, University of Ljubljana, Večna pot 113, Ljubljana 1000, Slovenia

^b Institute of Cell Biology, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, Ljubljana 1000, Slovenia

^c Visual Computing Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

ARTICLE INFO

Article history:

Received 8 March 2021

Revised 2 February 2022

Accepted 14 June 2022

Keywords:

Intracellular compartments

Reconstruction

Instance segmentation

Electron microscopy

Mitochondria

Golgi apparatus

Fusiform vesicles

Urothelium

Deep learning

ABSTRACT

Background and objectives: In recent years, electron microscopy is enabling the acquisition of volumetric data with resolving power to directly observe the ultrastructure of intracellular compartments. New insights and knowledge about cell processes that are offered by such data require a comprehensive analysis which is limited by the time-consuming manual segmentation and reconstruction methods.

Method: We present methods for automatic segmentation, reconstruction, and analysis of intracellular compartments from volumetric data obtained by the dual-beam electron microscopy. We specifically address segmentation of fusiform vesicles and the Golgi apparatus, reconstruction of mitochondria and fusiform vesicles, and morphological analysis of the reconstructed mitochondria.

Results and conclusion: Evaluation on the public UroCell dataset demonstrated high accuracy of the proposed methods for segmentation of fusiform vesicles and the Golgi apparatus, as well as for reconstruction of mitochondria and analysis of their shapes, while reconstruction of fusiform vesicles proved to be more challenging. We published an extension of the UroCell dataset with all of the data used in this work, to further contribute to research on automatic analysis of the ultrastructure of intracellular compartments.

© 2022 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Volumetric distribution of intracellular compartments and their temporospatial dynamics of interactions define the function and state of eukaryotic organisms [1]. The Golgi apparatus (GA), mitochondria, and endolysosomal compartments play the main role in the modification and sorting of membrane proteins, cell energy production, and degradation of endocytosed or autophagic materials in all cells [1]. Although typical functions are assigned to these organelles, there are differences in their distribution, organization and consequently, in the roles within different cell types. Moreover, there can be large differences between differentiation stages of the same cell types, and among normal cells and cells with dis-

eases (i.e., the structure of organelles may reflect pathological conditions) [2–5].

The urothelium is the epithelium of the urinary bladder and presents a good model to study intracellular compartments involved in the formation and polarized transport of membranes [6]. In differentiated urothelial cells, GA glycosylates transmembrane proteins (i.e., uroplakins), which are organized into 2D crystalline membrane domains (i.e., urothelial plaques) in the post-Golgi compartments are termed fusiform vesicles (FVs) [7–9]. FVs transport urothelial plaques to the apical plasma membrane of superficial urothelial cells, where they crucially contribute to the blood-urine barrier [10]. Plaques, endocytosed from the plasma membrane, or other compounds of the cell, can be degraded and removed by the endolysosomes if sufficient mitochondria-derived energy (i.e., ATP) is available [11]. It has been shown that many bladder diseases (e.g., urinary bladder cancer, cystitis, uroinfections), chemical agents, aging, or even relatively distant events (e.g., spinal cord injuries), affect urothelial functions, and that is reflected in the altered ultrastructure of the GA, mitochondria, and endolysosomes.

* Corresponding author.

E-mail addresses: manca.zerovnik-mekuc@fri.uni-lj.si (M. Žerovnik Mekuč), ciril.bohak@fri.uni-lj.si (C. Bohak), eb1690@student.uni-lj.si (E. Boneš), samo.hudoklin@mf.uni-lj.si (S. Hudoklin), rok.romih@mf.uni-lj.si (R. Romih), matija.marolt@fri.uni-lj.si (M. Marolt).

somes [12]. Thus, observing the ultrastructure of these compartments may lead to novel findings on how different factors affect the cells.

Electron microscopy (EM) is the only method with resolving power to directly observe the ultrastructure of intracellular compartments. In recent years, the availability of dual-beam microscopes (i.e., FIB-SEM) paved the way to acquire volumetric EM data on a large scale [13,14], but their (manual) segmentation and reconstruction, needed to study the ultrastructure of the compartments, is a major bottleneck, both in terms of time and expertise. Computer-assisted automatization of the process is expected to significantly facilitate and simplify analysis of volumetric electron microscopy data, however, currently available approaches are still not accurate enough.

We are addressing this bottleneck by proposing software methods for fully automatic segmentation and reconstruction of intracellular compartments from the FIB-SEM data. We refer to segmentation as the process of labeling each voxel of a volume with its corresponding semantic label, e.g., mitochondria, GA, fusiform vesicle and background (semantic segmentation), whereas we refer to reconstruction as the process of assigning organelle voxels to individual instances, thus separating them (instance segmentation).

This work is an extension of our previous work on segmentation of mitochondria and endolysosomes [15]. We present the following novel contributions: (1) methods for segmentation of FVs and the GA in FIB-SEM data, (2) reconstruction methods for FVs and mitochondria, and (3) methods for morphological analysis of mitochondria. Except for the methods for the reconstruction of mitochondria and segmentation of the GA, we are to our knowledge the first to present methods for the described tasks and we believe our work will enable further studies of the ultrastructure of cellular compartments. In addition, we also make public an extension of the UroCell dataset [15] by including all of the data used in this work.

2. Related work

We divide related work into three sections: segmentation, reconstruction and datasets.

2.1. Segmentation

Segmentation of volumetric data labels each voxel of the 3D volume with an appropriate label (e.g., mitochondria, endolysosome, background). It has become a fast-growing field with the success of convolutional deep neural networks, which can be applied to complex segmentation problems. The most widely adopted architectures in the field of volumetric segmentation of biomedical images such as the 3D U-Net [16], the V-Net [17], DeepMedic [18], and HighRes3DNet [19] are achieving high segmentation accuracy in different domains. Although general segmentation approaches can be successful, added expert knowledge about the target domain is usually as important as the choice of the segmentation network [20]. Lately Isensee et al. [21] presented the importance of parameter configuration. With the nnU-Net self-configuring method for deep learning-based biomedical image segmentation, they showed that with the right choice of parameters they can overcome most of the existing state-of-the-art methods on many public challenges.

In our work, we present a method for segmentation of FVs and the GA. We are aware of only a few works that address automatic segmentation of the GA, and none so far for FVs. Recently, Müller et al. [22] presented a method for segmentation of the GA from FIB-SEM data of the primary mouse β cell using the 3D U-Net architecture. The focus of the approach is on the reconstruction of

microtubule-organelle interaction in cells and not so much on segmentation accuracy, therefore they do not provide an evaluation of the segmentation algorithm. Another method for segmentation of multiple organelles, including the GA, was presented by Heinrich et al. [23]. Their method is based on several multi-channel 3D-U-Net architectures, that predict signed tanh boundary distances of binary labels for each organelle class. The distances are finally converted to binary labels of each target class. The authors present an extensive dataset and segmentation approaches for 35 different organelle classes, which will have a significant impact on the development of the field. The GA is segmented in the same fashion as all other organelles, but the authors do not evaluate its segmentation, since the test volumes do not include GA.

In the paper, we make the following contributions to segmentation of EM data. We present a new method for segmentation of the GA, which in contrast to the presented works does not require a precisely annotated training dataset. This is important, because such annotations are very time consuming to obtain due to the complex shape of the GA. We also quantitatively evaluate the segmentation accuracy of the proposed method on a dataset obtained from the urothelial tissue. We should note that the used dataset is challenging for segmentation of the GA, because it contains stacks of FVs that look very similar to the GA. Additionally, we present a method for segmentation of FVs, which to our knowledge is the first of its kind.

2.2. Reconstruction

Segmentation labels individual voxels with their respective class labels. With reconstruction, the individual compartment instances and their shapes are determined from the output of the segmentation process, thus enabling the counting of the compartments, measurement of their lengths, volumes, distribution, etc.

Works in automatic reconstruction of volumetric electron microscopy data are frequent in brain research, where the task, also referred to as connectomics, aims to reconstruct the full neuroanatomy, including synaptic contacts [24–28]. The reconstruction methods are not generally applicable to other domains, as they focus on the morphology of the target domain. Related to reconstruction, instance segmentation finds and extracts individual instances of objects in images and volumes. Many successful instance segmentation algorithms are derived from the MaskRCNN [29]. Even though there are examples of successful application of such algorithms to volumetric data (e.g., [30–32]), their reliance on bounding boxes is limiting with closely spaced objects or elongated curved objects, as the boxes may highly overlap. This is also the main reason that they are not used for instance segmentation of neurons.

In our work, we present methods for reconstruction of mitochondria and FVs in EM data. Segmentation of mitochondria has been addressed by many authors [33–37], a detailed overview has been presented in our previous paper [15]. Lately, several methods were proposed that also include reconstruction, where the main goal is to rectify locally incorrect segmentation results. Liu et al. [35] presented a method for automatic reconstruction of 3D mitochondria from automatic 2D segmentations. They use a multi-layer information fusion algorithm and in the end, discard all objects with length smaller than a selected threshold. Heinrich et al. [23] present a method to prevent over merging of instances in the segmentation output. After smoothing, hole filling, and size filtering, they create an over-segmentation of objects using a watershed algorithm which is followed by an agglomeration, where neighboring segments are merged together depending on the voxel values along the shared edges. Both mentioned works describe reconstruction of mitochondria, which is used to extract parameters for biomedical analysis (e.g., length, size, po-

sition...), but they are evaluated only voxel-wise. Such evaluation is valuable to show improvements in segmentation, however to show that the reconstruction and thus the calculated mitochondria parameters are reliable, it is necessary to evaluate the results with metrics that are used for object detection or instance segmentation. Namely, voxel-based metrics show little difference in cases where several neighboring mitochondria are falsely merged into one instance (and vice-versa), because such fusions consist of a negligible number of voxels in comparison to the size of mitochondria. Such errors will however affect the counting of instances, estimation of object sizes, etc. Recently Wei et al. [38] presented a large dataset for instance segmentation of mitochondria, where they showed that existing reconstruction methods are not very accurate when evaluated object-wise. The method with the best performance for instance segmentation of mitochondria on their dataset, is 3D U-Net, which predicts instance contours and binary masks, followed by a marker-controlled watershed (U3D-BC+MW), which we also evaluate on our data and show that our approach yields better results.

In this paper we present three contributions to reconstruction of EM data. We introduce a benchmark method for reconstruction of mitochondria based on the segmentation algorithm presented in Mekuč et al. [15]. The method consists of a set of post-processing steps that make use of topological properties of mitochondria [39] to reconstruct their shapes. We evaluate the approach with object-wise metrics. We also introduce novel methods for morphological analysis of mitochondria, by classifying their shape as classical, branched, and contacting. The introduction of such shape-based classification is important for biomedical analysis since experimental evidence demonstrate that mitochondrial ultrastructure affects mitochondrial and cellular function [40–42]. Both contributions are novel, as in the related work, authors usually calculate only the basic quantification parameters, such as counts, volume and surface area, and don't evaluate their accuracy [23,43]. Finally, we present and evaluate a method for reconstruction of FVs, which, to our knowledge, is the first of its kind. The proposed method is unsupervised and generic and can be applied to reconstruction of objects with similar topological properties.

2.3. Datasets

The evolution of the fields of automatic segmentation and reconstruction of intracellular compartments from volumetric microscopy data is highly dependent on the availability of annotated EM data.

Among the intracellular compartments that we are focusing on, several public volumetric datasets exist with annotated mitochondria [15,22,23,34,38,44–47]. For the GA, there are to our knowledge only two public volumetric FIB-SEM datasets [22,23], whereas no public datasets with FVs are currently available. All the datasets consist of binary class labels, except for the one presented by Wei et al. [38], which also provides labels for instance segmentation of mitochondria. Another relevant recently published dataset is the CEM500K dataset [48], which is a large unlabeled EM dataset beneficial for unsupervised pre-training of deep network features for segmentation of intracellular compartments.

With this paper, we make public an extension of the UroCell dataset [15], which contains annotations for mitochondria and endolysosomes. We extend it with manual annotations of FVs and the GA. We also add labels for instance segmentation of mitochondria and FVs, as well as labels for classification of mitochondria shape as classical, branched or contacting. This is to our knowledge the only public dataset obtained from the urothelial tissue, as well as the only one containing labels for instance segmentation

of FVs and classification of mitochondria instances based on their morphological properties.

3. Materials and methods

In this section, we describe the dataset and the proposed methods for segmentation and reconstruction of intracellular compartments. In Section 3.1 we first describe the dataset used for training and evaluation of the proposed methods. We present a method for segmenting FVs and the GA in Section 3.2, and the reconstruction methods for mitochondria and FVs in Section 3.3.

3.1. Data

The data that we use were acquired from tissue sample of the urinary bladder of male mice with a FIB-SEM dual-beam electron microscope. The obtained volume is of size $1560 \times 1366 \times 1180$ voxels, where the voxel dimensions are approximately $x = 16$ nm, $y = 16$ nm, and $z = 15$ nm. Details about the data acquisition are provided in Mekuč et al. [15]. The dataset used in this study consists of several sub-volumes ($256 \times 256 \times 256$ voxels) of the entire volume, cut out from different locations to make the dataset as diverse as possible.

This data is publicly available in the UroCell dataset [15]. Next to the raw data, it includes manual annotations of mitochondria and endolysosomes in five sub-volumes. There are from 15 to 85 mitochondria and approximately 10 endolysosomes in each of the sub-volumes. In this work, we extend the dataset with additional annotations, as described below. The enriched dataset is available on: <https://github.com/MancaZerovnikMekuc/UroCell> under the CC-BY-NC-SA 4.0 licence.¹ All the manual annotations were created with the open-source software Slicer3D [49]. In the end, they were revised by a cell biologist with expertise in urothelial biology. An example of a sub-volume showing raw data as well as manual annotations is shown in Fig. 1.

3.1.1. The extended dataset

In this work, we extend the dataset with manual annotations of FVs and the GA. We also include non-binary instance labels for mitochondria and FVs, and labels denoting branched and contacting mitochondria instances.

The annotations for FVs are given in four sub-volumes. There are 253 to 574 instances of FVs per sub-volume. For the GA, we provide precise annotations for two sub-volumes, which we use for evaluation of the proposed segmentation algorithm. For seven sub-volumes we provide an approximate manual annotation, where the region of the GA is roughly labeled. Every annotated sub-volume contains from 0 to 3 instances of the GA.

The labels for instanced, branched, and contacting mitochondria are provided for the five sub-volumes that also contain binary mitochondria labels. To label the instances, we inspected all of the existing binary components and divided or merged them, based on a thorough manual inspection of the underlying data. Labels for branched, and contacting mitochondria are based on the obtained instances and were defined with the help of a domain expert. We provide two sets of labels, one where each mitochondria instance is labeled as branched or non-branched and the other where each mitochondria instance is labeled as contacting or non-contacting. Mitochondria that are not in contacting and branched category are a part of classical category.

Classical mitochondria show “normal”, clearly outlined organelles typically found in cells. Branched category shows mitochondria with branched architecture. We call the third category

¹ <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode>.

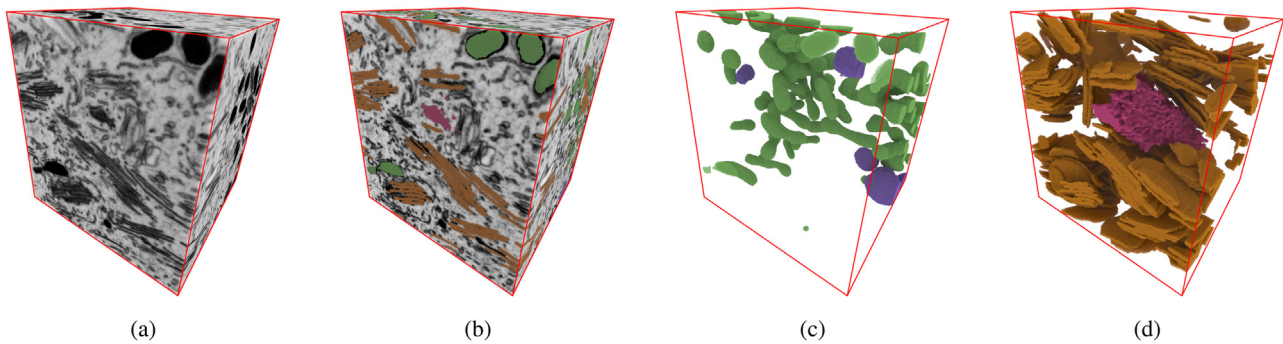


Fig. 1. A sub-volume from the UroCell dataset. (a) Raw FIB-SEM data. (b) Raw data plus manual annotations for mitochondria (green), endolysosomes (purple), FVs (orange), and the GA (pink). (c) Manual annotations of the mitochondria (green) and endolysosomes (purple). (d) Manual annotation of the FVs (orange) and precise manual annotation of the GA (pink).

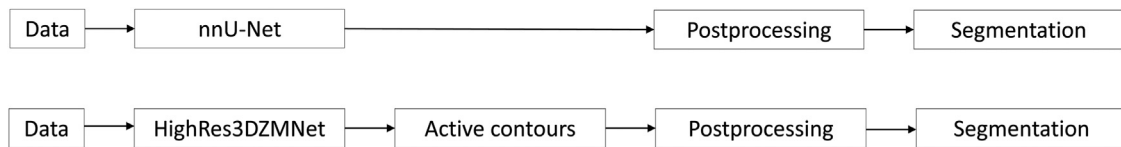


Fig. 2. Steps involved in the segmentation of FVs (upper diagram) and the GA (bottom diagram).

contacting mitochondria: such situations represent an intermediate state during mitochondrial fission or fusion, and the ability to sort out these mitochondria gives researchers the opportunity to assess the significance of this situation. We label the mitochondria as contacting if there is a visible narrow connection between the instances, but the category is borderline: they could represent functional contacts or be a technical artifact. Nevertheless, it is important that the interpretation is available to the researcher. The necessary certainty is going to be obtained through the accumulation of knowledge gained through advances in electron microscopic and complementary methods.

3.2. Segmentation

We present methods for segmentation of FVs and the GA (outlined in Fig. 2). Both methods use deep neural networks and different postprocessing steps to improve the segmentation accuracy.

3.2.1. Segmentation of fusiform vesicles

For segmentation of the FVs, we used the nnU-Net method described in Isensee et al. [21]. In the postprocessing stage, components smaller than $4 \times 10^5 \text{ nm}^3$ were removed (an average size of the FVs in the dataset is approximately $2 \times 10^7 \text{ nm}^3$). We also removed FVs that were detected over an area with curtaining artifacts, as these are usually false positives due to the artifacts.

3.2.2. Segmentation of the Golgi apparatus

The GA has a complex outline, due to its numerous cisternae and tubulo-vesicular network, which is difficult and time-consuming to accurately annotate manually - obtaining accurate ground truth for deep learning is thus difficult. We therefore propose a segmentation approach that relies on approximately annotated ground truth, where GA borders are only roughly traced; such annotations can be obtained much faster. An example of an approximate and precise annotation of a GA is shown in Fig. 3.

We train the HighRes3DZMNet network with approximate annotations. Highres3DZMNet is a deep neural network for segmentation of volumetric data, first presented in Mekuč et al. [15]. It consists of 19 layers with $3 \times 3 \times 3$ convolutions and one last $1 \times 1 \times 1$ convolution layer followed by a softmax. Convolutions

in the first seven layers are 1-dilated, in the next six layers they are 2-dilated and in the following six layers, they are 4-dilated. Between every two consecutive layers, except for the first and the last one, are residual connections. In the first layer, the network uses zero-mean convolutions, which makes it more robust to varying brightness levels in different patches of the input volume. Because the network is trained with approximate annotations, it will output volumes with regions that likely contain the GA, however GA borders will not be accurately segmented. To refine the segmentation, we use the active contours algorithm [50]. The active contours, also called snakes, are an iterative region-growing process that pushes the initially given object estimate to the object boundaries by minimizing an energy function. In our case, the initial estimates are segmentations obtained by the deep network. We use a 2D active contours algorithm [51], which we run on every slice of the analyzed volume along all three dimensions. This results in three different output volumes, which we merge using disjunction (a voxel is considered as part of the GA if it is labeled as GA in at least one of the three volumes).

The GA is morphologically similar to FVs that are very common in our dataset. Thus, false positives may occur as the vesicles get marked as a GA by the segmentation network. We therefore added an additional postprocessing step that removes such false positives. Because FVs are usually smaller in size, this step simply removes all detected GAs smaller than $0.19 \mu\text{m}^3$ which is 80% of the smallest GA in our ground truth.

With the described approach we obtain a precise segmentation of the GA even though we only rely on approximate ground truth for training, while the process is still completely automatic.

3.3. Reconstruction

Each segmentation network outputs a volume, where each voxel is labeled with its corresponding class (or as background). To count the organelles in a class and to make quantitative analyses of their sizes, shapes and distribution, we need to reconstruct individual instances of these organelles. This is trivial, if their voxels are already separated; however mitochondria and FVs are often found in groups with their membranes touching, thus forming large sets of connected voxels, which must be divided into individual instances.

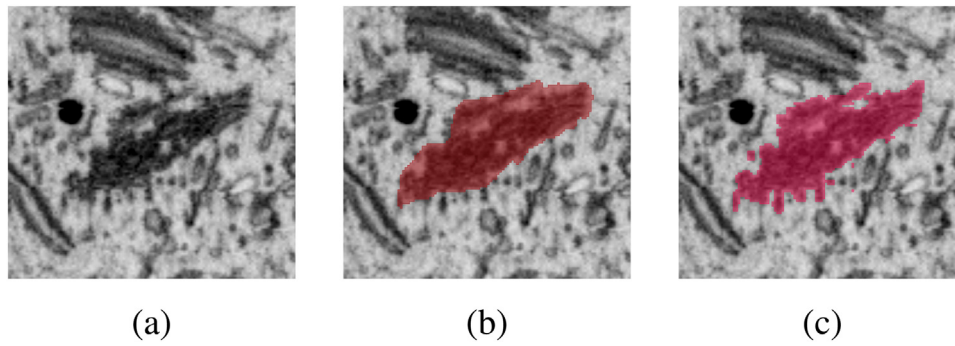


Fig. 3. A 2D slice containing the GA. (a) Raw data. (b) Approximate manual annotation. (c) Precise manual annotation.

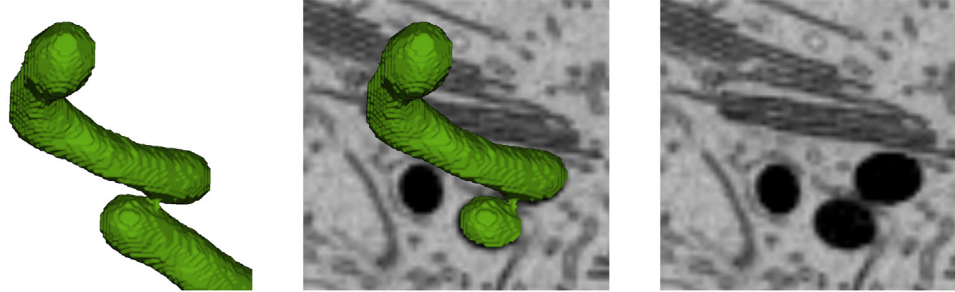


Fig. 4. Two mitochondria, classified as one connected set of voxels (left). Analysis of EM data (right) shows that they are only touching and are thus two separate instances.

This is the goal of reconstruction - extracting the 3D shape of each organelle instance. We propose two algorithms for this purpose, both making use of the domain knowledge on the shape of individual organelles. The algorithms are designed for reconstructing mitochondria and FVs. Although we also segment the GA and endolysosomes, these are very rarely touching in our dataset, so their reconstruction is simple. In addition, we present methods for exploiting the reconstructed information to classify mitochondria shape into classical, branched and contacting.

3.3.1. Reconstructing mitochondria

Mitochondria may often touch, thus forming connected sets of voxels. The goal of reconstruction is to separate such instances. An example is shown in Fig. 4.

The main goal of reconstruction is to separate closely-spaced mitochondria, which are falsely connected across membranes, because the deep network labeled the voxels between them as mitochondria (instead of background). We use the active contours algorithm to shrink the labeled components towards mitochondria membranes and thus separate the mitochondria. We use the edge-based model for active contours, which is similar to geodesic active contours [52]. We set the maximum number of iterations of evolving contours to 3 and set the smoothing factor and contraction bias to 0. We run the algorithm along each dimension of the 3D volume, which we merge as already described for segmentation of the GA. After applying the process of active contours, we perform noise removal by removing all components labeled as mitochondria and smaller than $4.7 \times 10^6 \text{ nm}^3$, which is approx. 50% of the size of the smallest mitochondrion in the ground truth. In this way, we remove small noisy chunks sometimes attached to mitochondria as the result of poor segmentation (often near the cell membrane), and now disconnected by our reconstruction method.

3.3.2. Reconstructing fusiform vesicles

FVs densely populate the urothelial tissue (see Fig. 1d). Their shape can be described as a pancake in 3D space and they are usually found in stacks, which makes it difficult to distinguish be-

tween individual instances. Successful segmentation will result in blobs of voxels classified as FVs, however to actually analyze their counts, sizes and interactions, we need a reconstruction method to divide the blobs into individual vesicle instances. An example of such stacks and blobs is shown in Fig. 5.

Because we are dealing with thin objects, placed very close together, existing methods for volumetric instance segmentation, such as the MaskRCNN [29], do not yield good results, as the bounding boxes overlap too much. We therefore present a new unsupervised reconstruction method that can be used as a benchmark for the problem. It is based on fitting thin disks to the vesicle stacks, output by the segmentation algorithm. As fitting multiple 3D shapes in 3D space can be a challenging optimization problem, due to the many free parameters, we add steps that limit the number and ranges of parameters. We outline the method in Algorithm 1 and describe it in more detail below.

```

input : Segmentation of FVs denoted by vol and its
         corresponding FIB-SEM volume denoted by volFIB
output: A set of divided FVs, denoted by  $\mathcal{V}$ 
1  $\mathcal{V} \leftarrow \emptyset$ ;
2 foreach connected component ci in vol do
3   fibi, rotinfo  $\leftarrow$  CutFromFIB(ci, volFIB);
4   Clrot  $\leftarrow$  Rotate(ci, fibi);
5   disks  $\leftarrow$  FitDisks(Clrot);
6   disksrotated  $\leftarrow$  RotateBack(disks, rotinfo);
7   ciclassified  $\leftarrow$  NearestValue(disksrotated, ci);
8    $\mathcal{V} \leftarrow \mathcal{V} \cup \text{ciclassified}$ ;
9 end
10 return  $\mathcal{V}$ ;

```

Algorithm 1: Reconstruction of FVs.

The reconstruction method operates on the output of the algorithm for segmentation of FVs, as well as on the corresponding original FIB-SEM volume. We first find all connected components (vesicle stacks) in the input volume. Each stack is first ro-

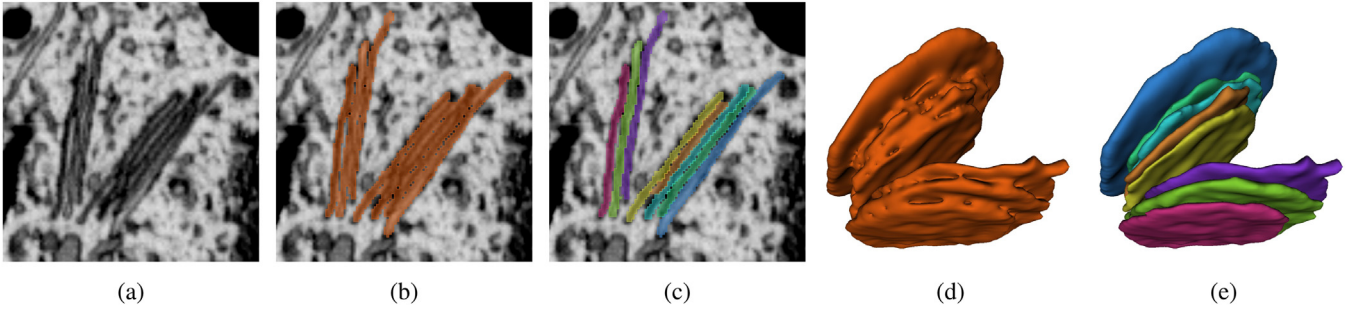


Fig. 5. Two stacks of FVs. (a) A 2D slice from the FIB-SEM data showing the two stacks. (b) Manually labeled ground truth for binary segmentation. (c) Instanced ground truth. (d) Binary ground truth shown in 3D. (e) Instanced ground truth shown in 3D.

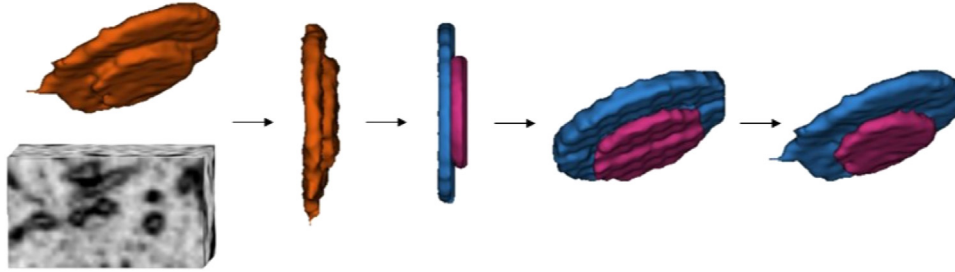


Fig. 6. Reconstruction of FVs. The steps are illustrated left-to-right as follows: first the vesicle stack segmented from the FIB-SEM data is shown. The stack is then rotated to be parallel to the sagittal plane. Two disks were fitted with the proposed optimization approach, then rotated back and finally merged with the original segmentation, where a separate label is assigned to each vesicle instance.

tated so that the vesicles are parallel to the sagittal plane (line 4 of Algorithm 1). To determine the rotation angle and axis, we analyze the cross section of the stack within the original FIB-SEM volume. We estimate stack orientation by combining edge detection with the Hough transform to detect lines parallel with vesicles in the stack. According to the average line rotation, we rotate the stack to position it parallel to the sagittal plane. An example is shown as the first step in Fig. 6.

In the next step, we use an optimization approach to fit a set of disks, representing individual vesicles, to the rotated stack. The disks represent a rough approximation of the vesicle shape, we use them because their shape can be described by only one parameter (radius), in contrast to more flexible representations such as parametric surfaces. Together with their 3D position, each disk is defined by 6 varying parameters and one fixed parameter (thickness) listed below. As we start the optimization with a stack that is parallel to the sagittal plane, we can make a good initial estimate of parameter values for each disk. We can also limit their range of values, in order to reduce the optimization search space, as we already know their approximate orientation and have knowledge of the typical sizes of FVs. We use the following parameters and ranges (in voxels):

- radius $r \in [\frac{0.4 \cdot s_y}{2}, \min(\frac{s_y}{2} + 5, 30)]$,
- coronal rotation $cro \in [cro_{curr} - 10, cro_{curr} + 10]$,
- axial rotation $aro \in [aro_{curr} - 10, aro_{curr} + 10]$,
- position $x \in [0, s_x]$,
- position $y \in [s_y - 10, s_y + 10]$,
- position $z \in [0, s_z]$.
- thickness $h = 4$,

s_x , s_y and s_z denote the size of the bounding box of the vesicle stack, and cro_{curr} , aro_{curr} the coronal and axial rotation of the current optimization iteration. Rotations are initially set to zero, while other parameters are set to the middle of their ranges. Thickness

is constant and does not vary, its value was determined from observation of the vesicles in our data.

We perform optimization with a pattern search algorithm minimizing an objective function with bound constraints [53]. To estimate the number of vesicles in the stack, we iteratively add new disks and run the optimization algorithm, while the objective decreases by more than a threshold (we selected the threshold value of 0.02).

We define the objective function as follows:

$$f(c, v, n) = 0.7 (1 - dsc(c, v, n)) + 0.3 ovl(v, n); \quad (1)$$

where c is the analyzed component, v is a volume containing the already fitted disks and n the new disk that is being added in the current iteration. The first part of the objective function measures the Dice similarity coefficient (defined in Section 4.2) between the component and the volume containing the fitted disks - the better the fit, the larger the value. As disks should not overlap, the second part of the objective function ($ovl(v, n)$) returns the percent of overlapping disk voxels thus penalizing the overlap. An example of the output of this step is shown in Fig. 6.

After the disks are fitted, we rotate them back to their initial orientation. In the end, we assign instance labels to the voxels by assigning each voxel to its nearest disk (lines 6 and 7 of the Algorithm 1 and the last two steps in Fig. 6).

3.4. Shape classification

Reconstruction enables researchers to obtain detailed information on the distribution, organization, sizes and shapes of cellular compartments. This in turn enables the study of the differences in compartments between different cell types, between differentiation stages of the same cell types, and among normal cells and cells with diseases (i.e., the structure of the compartments may reflect pathological conditions). To this end, we present two approaches to classify mitochondria according to their shape as classical, branched and contacting.

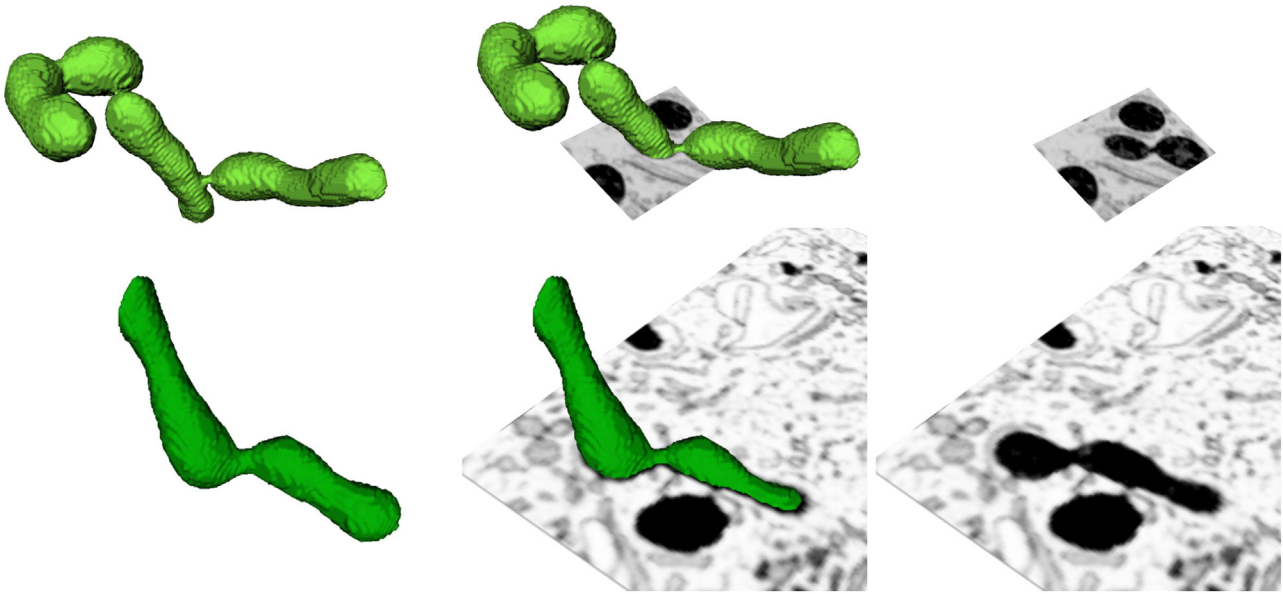


Fig. 7. Two examples of contacting mitochondria. The upper example has two contacting sections, the lower has one contacting section. For each mitochondria, we show the underlying microscopy data.

3.4.1. Detection of contacting mitochondria

Contacting mitochondria contain regions, where they are significantly narrower with respect to their overall shape. Examples of such objects are shown in Fig. 7.

The contacting mitochondria are detected within the set of reconstructed mitochondria instances obtained by the algorithm described in Section 3.3.1. The main idea of the proposed approach is to calculate the skeleton of each instance and then detect the narrowed parts by observing the distances of voxels within the instance to the skeleton. The approach is outlined in Algorithm 2.

input : Set of reconstructed mitochondria instances denoted by \mathcal{M} output : Set of contacting mitochondria instances denoted by \mathcal{M}_c	<pre> 1 $\mathcal{M}_c \leftarrow \emptyset$; 2 foreach instance i in \mathcal{M} do 3 $i \leftarrow \text{FillHoles}(i)$; 4 $\text{skel} \leftarrow \text{Skeleton}(i)$; 5 $\text{skeldist} \leftarrow \text{Distance}(i, \text{skel})$; 6 $\text{skeldist}(\text{skeldist} < \theta_1) = 0$; 7 $n \leftarrow \text{CountComponents}(\text{skeldist})$; 8 if $n > 1$ then 9 $\mathcal{M}_c \leftarrow \mathcal{M}_c \cup i$; 10 end 11 end 12 return \mathcal{M}_c; </pre>
--	--

Algorithm 2: Algorithm for detection of contacting mitochondria.

The method iterates through all the reconstructed mitochondria instances. Holes within each instance are first filled by a simple hole filling algorithm, to avoid errors in further steps. To obtain the center of the instance, its 3D medial skeleton is calculated with the homotopic thinning algorithm [54]. In the next step, the shortest distance to the skeleton is calculated for each instance voxel, thus adding a measure of thickness to the instance and resulting in an instance volume skeldist . All the voxels in skeldist with thickness lower than a selected threshold θ_1 are removed in the next

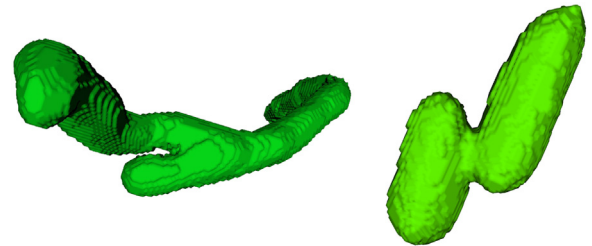


Fig. 8. Two examples of branched mitochondria.

step. The parameter θ_1 controls the thickness margin at which we denote a mitochondrion as contacting. In the final step, we count the components in the skeldist ; if more than one is found, the instance is classified as contacting. We do not count components smaller than 5 voxels, which sometimes appear at the ends of the instances, as they do not represent the narrowings that we are interested in.

3.4.2. Detection of branched mitochondria

Mitochondrial fusion and division gives rise to branched mitochondria instances, as well as entire mitochondria networks. Examples of branched objects are shown in Fig. 8. To automatically detect branched mitochondria, we again observe the reconstructed instances obtained by the algorithm described in Section 3.3.1. As with contacting mitochondria, we also observe 3D mitochondria skeletons, which are transformed into graphs. If these graphs contain multiple edges, the instance is labeled as branched. The approach is outlined in Algorithm 3.

The method iterates through all the reconstructed mitochondria instances. Holes within each instance are first filled by a simple hole filling algorithm, to avoid errors in further steps. The 3D skeleton is extracted with the same method as in the previous algorithm for detection of contacting mitochondria. The skeleton is then transformed to a graph [55], where branches smaller than a threshold θ_2 voxels are removed, as they likely represent small protrusions or noisy instance boundaries. In the final step, we add the instance labeled as branched if there are more than two edges in the obtained graph.

```

input : Set of reconstructed mitochondria instances denoted
        by  $\mathcal{M}$ 
output: Set of branched mitochondria instances denoted by
         $\mathcal{M}_b$ 
1  $\mathcal{M}_b \leftarrow \emptyset$ ;
2 foreach instance  $i$  in  $\mathcal{M}$  do
3    $i \leftarrow \text{FillHoles}(i)$ ;
4    $\text{skel} \leftarrow \text{Skeleton}(i)$ ;
5    $\text{graph} \leftarrow \text{SkeletonToGraph}(\text{skel}, \theta_2)$ ;
6    $n \leftarrow \text{CountEdges}(\text{graph})$ ;
7   if  $n > 2$  then
8      $\mathcal{M}_b \leftarrow \mathcal{M}_b \cup i$ ;
9   end
10 end
11 return  $\mathcal{M}_b$ ;

```

Algorithm 3: Algorithm for detection of branched mitochondria.

4. Experiments and results

4.1. Implementation details

For the segmentation of FV with nnU-Net we ran 4-fold cross-validation for all the proposed architectures. For the prediction, we used automatically selected 3D U-Net cascade. Predictions were made for each fold separately. We used the publicly available code [21].

HighRes3DZMNet for the segmentation of GA was implemented with the Nifty-Net framework [56]. We ran 30,000 iterations of training with a learning rate of 0.001 using the Adam optimizer. For both approaches, we used L2 regularisation with λ equal to 10^{-5} . The reconstruction algorithms are implemented in Matlab. The thresholds used by the shape classification methods were manually set based on observations of typical sizes of the analyzed compartments. Their values (in voxels) were set as: θ_1 to 4.25 and θ_2 to 2. When applied to volumes of different resolution, these should be changed accordingly. Code is available at <https://github.com/MancaZerovnikMekuc/HighRes3DZMNet> and <https://github.com/MancaZerovnikMekuc/Reconstruction>.

For the mitochondria reconstruction, we evaluated the U3D-BC+MW method presented in Wei et al. [38] to compare the methods on our data. Here we also ran 30,000 iterations of training. For the watershed algorithm, we used the same parameters as proposed for the MitoEM dataset.

4.2. Evaluation metrics

To quantitatively evaluate the results, we use the well-established metrics from the field of semantic segmentation and object detection: the Dice similarity coefficient (DSC) also called the F1 score (F1), intersection over union (IOU) also called the Jaccard similarity coefficient (JAC), true positive rate (TPR) also called recall (REC) or sensitivity, true negative rate (TNR) also called specificity, and precision (PRE). They are defined as follows:

$$\text{DSC} = \text{F1} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (2)$$

$$\text{IOU} = \text{JAC} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (3)$$

$$\text{TPR} = \text{REC} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

Table 1

Results for segmentation of FVs.

Method	DSC	TPR	TNR
nnU-Net [21]	0.757 \pm 0.03	0.701 \pm 0.05	0.979 \pm 0.00
Final segmentation	0.763 \pm 0.03	0.707 \pm 0.05	0.979 \pm 0.00

$$\text{TNR} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (5)$$

$$\text{PRE} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

In the upper equations TP (true positives) is the number of elements that are correctly classified as the target class, TN (true negatives) is the number of elements that are correctly classified as the background class, FP (false positives) is the number of elements which are falsely classified as the target class, and FN (false negatives) is the number of elements which are falsely classified as the background class. The highest possible value for all the metrics is 1, which we would get if all the elements would be classified correctly. Depending on the evaluated task, those elements can be voxels or individual compartment instances.

For evaluation of segmentation, we measure the DSC, TNR, and TPR. All the metrics are measured voxel-wise. DSC gives us information on the similarity between the ground truth annotations and the output of the segmentation. TPR measures the proportion of voxels that are correctly classified as the target class, and TNR measures the proportion of voxels that are correctly classified as the background class.

To evaluate reconstruction, where we extract individual compartment instances, we measure object detection accuracy. First, we determine the number of TP, TN, FP, and FN instances. An instance in the result is considered as a TP if it overlaps with an instance in the ground truth and if this overlapping, which is measured by an IOU metric voxel-wise, is higher than a selected threshold. If we have multiple instances for one ground truth object, the one with the highest IOU is considered as the TP and all the others are counted as FP. From the instance-wise TP, TN, FP, and FN, we calculate instance-wise PRE, REC, F1. PRE measures the proportion of correctly classified instance among all of the detected instance, REC measures the proportion of detected instance among all the ground truth instances, and F1 is the harmonic mean of PRE and REC. We also calculated the Panoptic Quality (PQ) metric [57], which aggregates the semantic and instance segmentation quality in one metric. It is defined as follows:

$$\text{PQ} = \frac{\sum_{(p,g) \in \text{TP}} \text{IOU}(p,g)}{\text{TP} + \frac{1}{2}\text{FP} + \frac{1}{2}\text{FN}} \quad (7)$$

In the upper equation the TP, FP, and FN are numbers calculated instance-wise with an IOU threshold of 0.5. In the numerator, p is a true positive object from the prediction and g stands for its corresponding object in the ground truth.

4.3. Results

In the following subsections, we present the qualitative and quantitative results of the proposed approach. It is important to look at both aspects, because the numbers do not always accurately describe the quality or usability of the results.

4.3.1. Segmentation results

Table 1 shows results for the segmentation of FVs. They were obtained by 4-fold cross-validation, so that each sub-volume from

Table 2
Results for segmentation of the GA.

Method	DSC	TPR	TNR
HighRes3DZMNet [15]	0.515	0.861	0.983
Active contours	0.501	0.876	0.982
Final segmentation	0.926	0.885	0.999

Table 3
Results for reconstruction of mitochondria.

Method	TP	FP	FN	PRE	REC	F1	PQ
U3D-BC+MW [38]	159	156	18	0.505	0.898	0.323	0.61
HighRes3DZMNet [15]	150	1142	27	0.116	0.847	0.204	0.19
Noise removal	151	25	26	0.858	0.853	0.856	0.81
Final reconstruction	157	22	20	0.877	0.887	0.882	0.82

the dataset was once used as a test volume. The table shows averages of all four runs, with and without postprocessing. We also visualize the segmentation of the sub-volumes with the highest and the lowest DSC scores and the corresponding ground truth in Fig. 10.

Table 2 shows results for segmentation of the GA. We show evaluation for each step of the proposed method. We trained the network on five approximately annotated sub-volumes and evaluated it on two sub-volumes where precise ground truth annotations were available (these were not included in training). One of the test subvolumes is also visualized in Fig. 11.

We also evaluated the time needed for annotation of the GA. As noted previously, exact annotation is extremely time consuming, as the shape of the GA is complex, so we based our segmentation algorithm on approximate annotations. An experienced annotator needed approximately 2.5 hours for precise annotation and 0.5 hours for approximate annotation of one GA, resulting in a 5-fold speed up for making the ground truth annotations.

4.3.2. Reconstruction results

Results for the reconstruction of mitochondria, based on instance-wise metrics, are shown in Table 3. They were calculated with an IOU threshold of 0.75 (the values remain very similar for other threshold values and the F1 measure only falls to 0.7 for the threshold value of 0.9). Reconstruction is based on the segmentation algorithm presented in Mekuč et al. [15], which we also evaluated; results are shown in the second line of the table. We evaluated the approach on all sub-volumes that contain instance labels. We summed the TP, FP, and FN for all sub-volumes and calculated the measures in the table from the summed metrics. As mitochondria touching the borders of a sub-volume may only be partially included in the sub-volume and may thus be incorrectly treated by the proposed approach, we removed the border objects smaller than 10,000 voxels from evaluation. A visualization of reconstruction on a 2D slice is shown in Fig. 9. With the same procedure, we evaluated the output of the U3D-BC+MW method on our dataset. The results are shown in the first line of the table. We can see that our approach outperforms the U3D-BC+MW method, especially because of its higher precision.

For reconstruction of FVs, we used the same instance-wise evaluation metrics. They were calculated from TP, FP, and FN values summed over all four evaluated sub-volumes. Results are presented in Table 4. In the table, we show the results calculated with IOU thresholds of 0.3 and 0.5. We compare the results with the output of the segmentation algorithm, where we label each connected component as a separate object. Results of the proposed algorithm are shown in the last line of both parts of the table. Visualization of reconstruction on one of the sub-volumes is shown in Fig. 12.

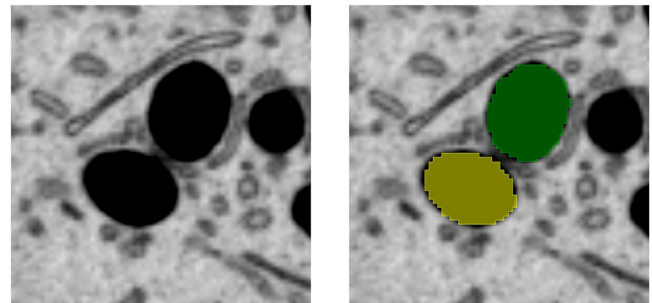


Fig. 9. A slice showing two mitochondria (green and yellow) that touched and were thus connected in the output of the segmentation (left). They were divided by the proposed reconstruction algorithm (right).

4.3.3. Shape classification

For the detection of contacting and branched mitochondria, we use the same evaluation metrics as for the reconstruction algorithm. Results are shown in Table 5. Again we used the IOU threshold of 0.75. Examples are visualized in Figs. 13 and 14.

4.3.4. Reconstruction of the entire volume

To demonstrate the capability of our approach, we ran the segmentation and reconstruction algorithms on the whole volume with resolution: $1280 \times 1024 \times 1024$ voxels. The results are visualized in Fig. 15, where one can see the dense packing of intracellular compartments within a differentiated superficial urothelial cell.

5. Discussion

In the following subsections, we discuss the presented results.

5.1. Segmentation

We presented two approaches for the segmentation of intracellular compartments, which were to our knowledge not yet evaluated on the FIB-SEM data. We also make public manual annotations for both classes as part of the UroCell dataset.

Our method for the segmentation of FVs achieved the average DSC score of 0.763. We can explain this value better by looking at the qualitative results. The method segmented the vesicles correctly most of the time (Fig. 10(b)), however, results were less satisfying for one sub-volume cut from the edge of the cell (Fig. 10(d)). This sub-volume contains more artifacts, as well as the cell membrane, which is sometimes very similar to the FVs and is thus incorrectly segmented (Fig. 16). Even though the method itself doesn't present much technical novelty, we pose new challenges to the field by providing benchmark results and analysis of the method's drawbacks.

The GA was successfully segmented with a high DSC score (0.926). In the result table, we can see that the DSC is lower after the second step of the method (active contours). The reason lies in the false positives which are usually enlarged by the active contours, but are later eliminated in the postprocessing step. The presented method didn't miss any of the GA in the test set, except for one on the border of the sub-volume, which was removed, because its part inside the sub-volume was too small. Several false positives were caused by stacks of FVs labeled as the GA and empty space in the middle of the GA, which was not labeled as empty space. The number of GA in the test set was relatively low, so we also ran the segmentation on the entire volume, which qualitatively showed that the method successfully segmented the GA also in other sub-volumes.

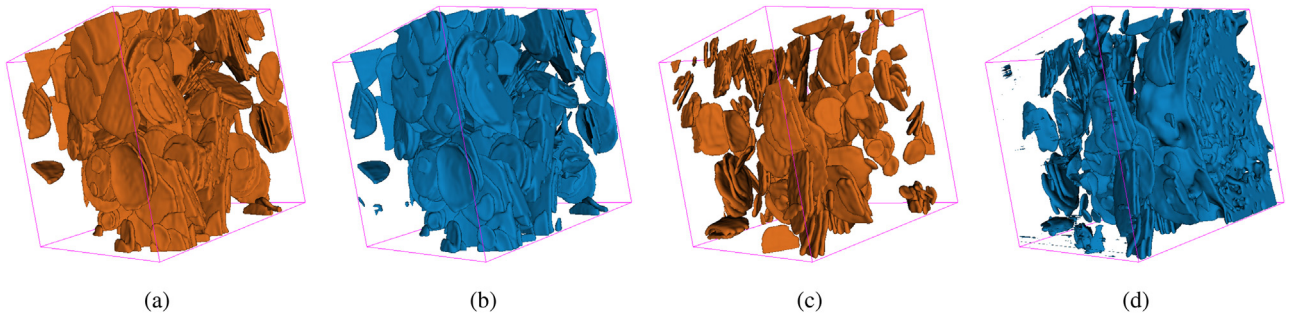


Fig. 10. Segmentation of FVs. (a) Ground truth of the sub-volume with the highest DSC. (b) Segmentation of the sub-volume with the highest DSC. (c) Ground truth of the sub-volume with the lowest DSC. (d) Segmentation of the sub-volume with the lowest DSC.

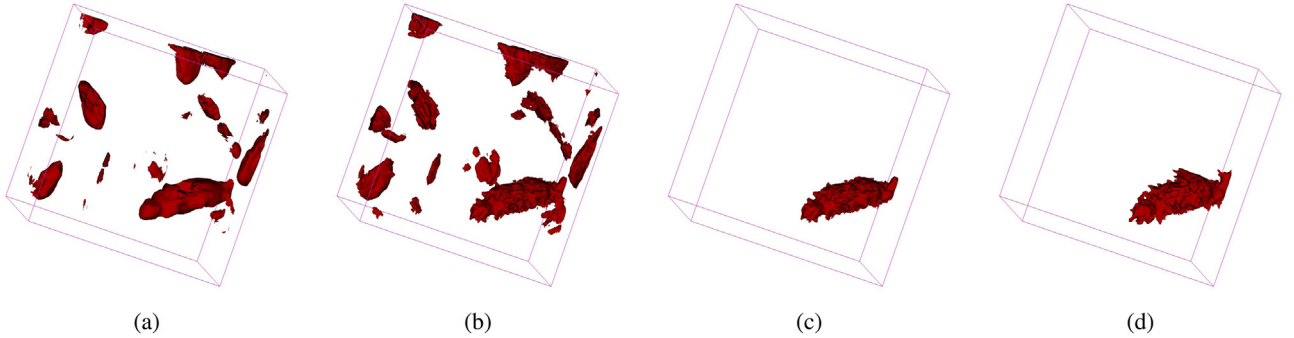


Fig. 11. Segmentation of the GA on one of the test sub-volumes. (a) Output of the HighRes3dZMNet. (b) Segmentation after applying active contours. (c) Final segmentation after postprocessing. (d) Ground truth.

Table 4
Results for the reconstruction of FVs.

Method	TP	FP	FN	PRE	REC	F1	PQ
IOU thr. = 0.3 Segmentation	183	391	1398	0.319	0.116	0.170	–
Reconstruction	250	295	1331	0.459	0.158	0.235	–
IOU thr. = 0.5 Segmentation	120	454	1461	0.209	0.076	0.111	0.085
Reconstruction	140	405	1441	0.257	0.089	0.132	0.088

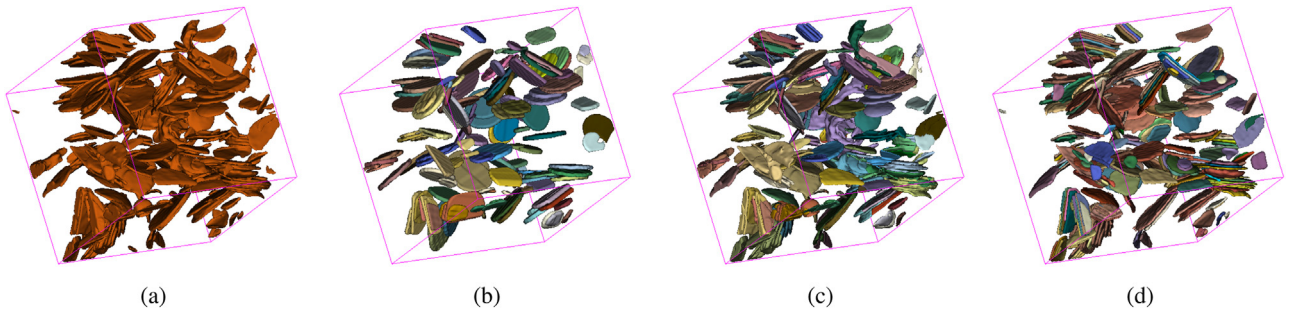


Fig. 12. Reconstruction of FVs. (a) Segmentation. (b) Disks fitted to the segmentation. Each disk has its own label. (c) The final output from the algorithm. (d) Ground truth. (a-d) Different colours present different labels. The colours don't match among the different sub-volumes. The only important thing is that a separate instance is of a different colour.

Table 5
Results for detection of contacting and branched mitochondria.

Method	TP	FP	FN	PRE	REC	F1
Contacting detection	16	6	13	0.727	0.552	0.627
Branched detection	11	4	11	0.733	0.500	0.595

With our approach, we also showed that we can get accurate segmentation results with approximate ground truth labels, thus significantly reducing the time needed for annotation (5-fold). We believe that a similar approach is also useful for other segmen-

tation problems, where objects with complex border needs to be segmented.

5.2. Reconstruction

Reconstruction of mitochondria consists of a few steps. Results in Table 3 show that the noise removal step removes many false positives and that the further step of dividing the mitochondria instances successfully divides most of the touching mitochondria. There are still some cases where division does not succeed, which happens if instances touch along a larger surface, so that the active

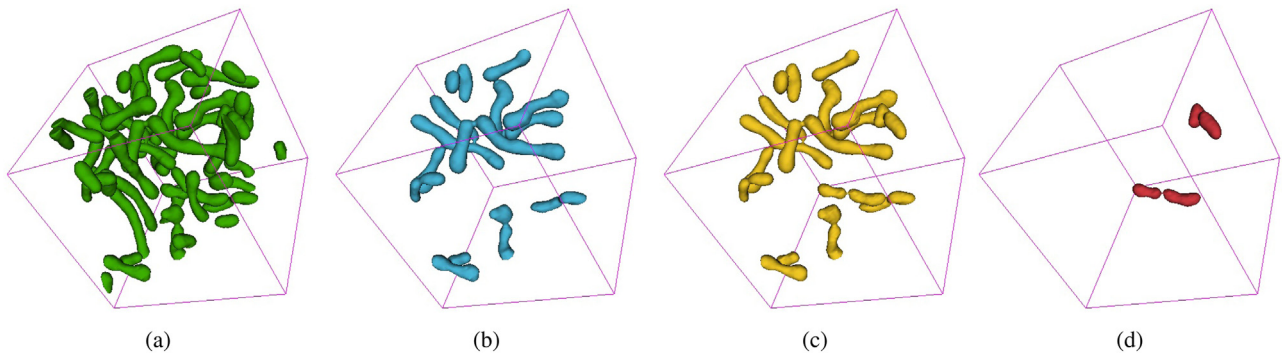


Fig. 13. Detection of contacting mitochondria. (a) Results of reconstruction. (b) The output of detection of contacting objects. (c) Ground truth. (d) Differences between the ground truth and the result showing two missing false negatives in red.

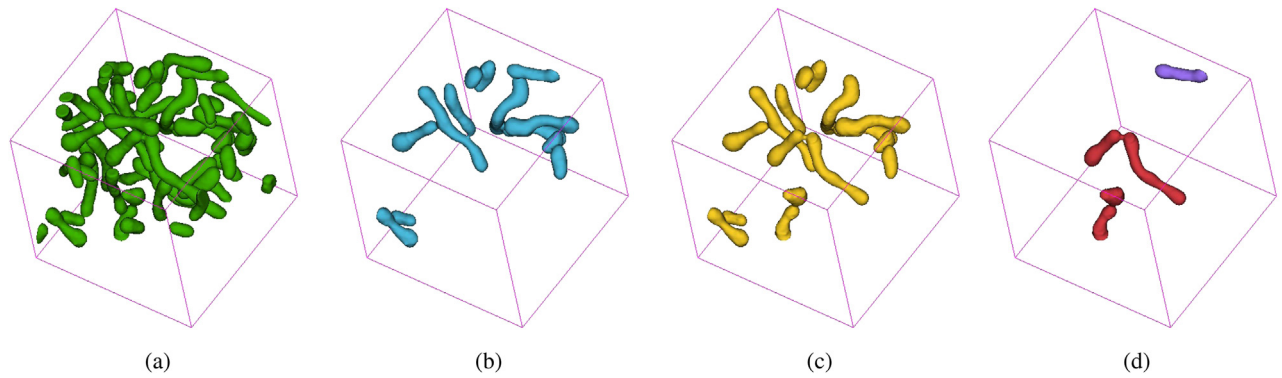


Fig. 14. Detection of branched mitochondria. (a) Results of reconstruction. (b) The output of detection of branched objects. (c) Ground truth. (d) Differences between the ground truth and the result. False negatives are labeled with red and false positives with purple.

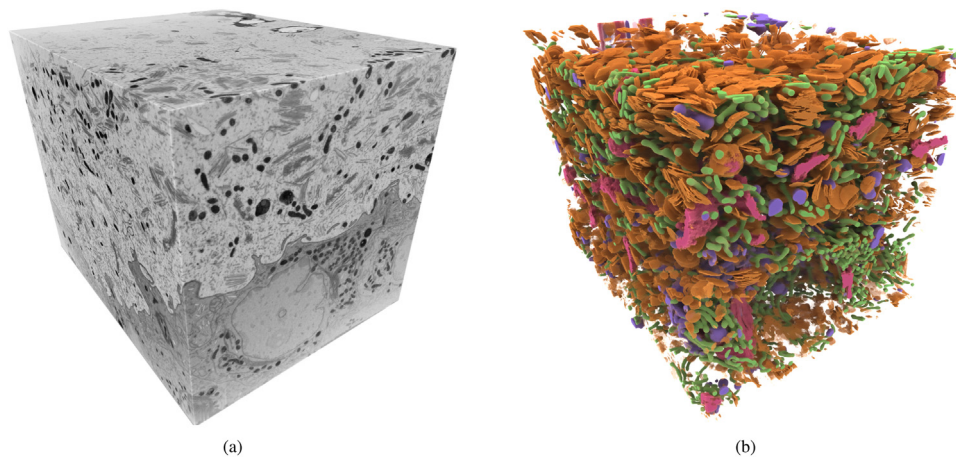


Fig. 15. Segmentation of the entire volume shows how densely populated the cells are with the intracellular compartments. (a) shows the raw microscopy data, and (b) shows the reconstruction with the proposed approach: mitochondria (green), endolysosomes (purple), FVs (orange), and the GA (pink).

contours method does not separate them. An example is shown in Fig. 17.

Reconstruction of FVs is a harder problem, especially because the vesicles are very thin and close together, which makes it hard to distinguish the individual instances, even to biomedical experts.

The results in Table 4 show a low F1 measure, although the increased number of true positives indicate that some stacks were successfully separated into vesicle instances. We can also observe this in the qualitative results visualized in Fig. 10. Analysis shows that simple stacks are usually well resolved (an example is shown

in Fig. 6), but the method suffers on the more complex ones, mainly because of the narrow parametrization and oversimplification of using disks as vesicle models. The vesicles are often curved, so disks do not cover their volume well. Vesicles may also have different orientations in a stack and if several stacks are touching, they will not be correctly resolved, yielding many false negatives. An example of multiple touching stacks is shown in Fig. 18.

Another problem arises with small vesicles, as the objective function will prefer fitting a disk to several vesicles, instead of a single smaller one. We visualize the problem in Fig. 19, where we

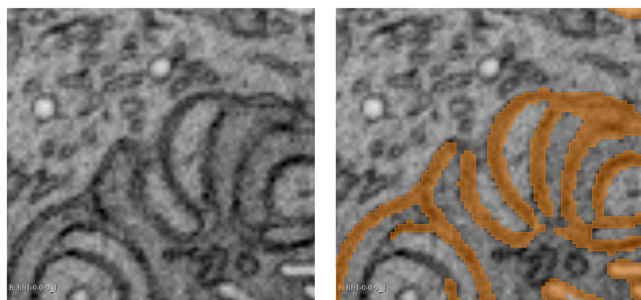


Fig. 16. Incorrect segmentation of the cell membrane (left) as FVs.



Fig. 17. An example of two mitochondria which are erroneously not divided.

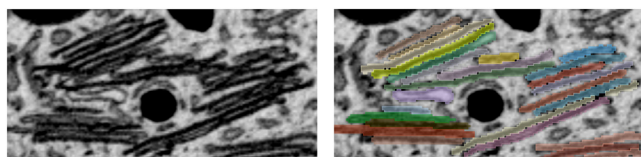


Fig. 18. Multiple stacks of FVs touching each other. Raw data is shown on the left, and the ground truth on the right.

can see that the red disk was fitted over two vesicles, which has a domino effect on the fitting of additional disks. Overall, the proposed method can be viewed as a baseline with many opportunities for improvements in the future.

5.3. Shape classification

We presented two algorithms for detecting contacting and branched mitochondria. The algorithms both start with the reconstructed mitochondria instances as their input, so their results are limited by the accuracy of reconstruction (and segmentation).

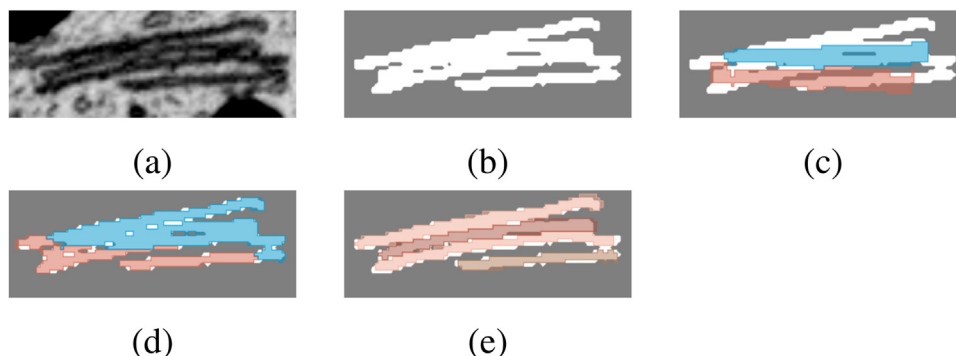


Fig. 19. Errors in reconstruction of FVs. (a) Raw data. (b) Output from the segmentation algorithm. (c) Disks fitted to the component from the segmentation step. (d) Final (incorrect) output of our approach. (e) Ground truth - four instances of FVs.

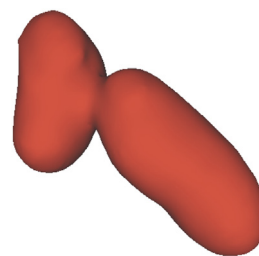


Fig. 20. An incorrectly labeled contacting mitochondria.

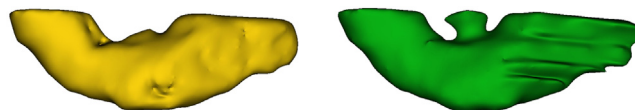


Fig. 21. A branched mitochondria (right) that was not correctly segmented (left) and was thus incorrectly labeled as non-branched.

The detection of contacting mitochondria is affected by errors in the input data, as erroneous touching or divisions of mitochondria are usually narrow. This can be observed in the lower two red mitochondria in Fig. 13(d), which have a narrow connection in the ground truth but not in the input volume of the detection algorithm. The method itself is sensitive to cases where the contacting section is quite thick and is thus not labeled as such. This may be observed in Fig. 13(d) - the upper red false negative object. We show an enlarged image of it in Fig. 20.

The detection of branched mitochondria is mostly functioning correctly. False positives and false negatives are usually caused by the problems in the input data: mitochondria which were not divided by the division algorithm and are therefore branched are detected as false positives (an example is the purple object in Fig. 14(d)), objects with incomplete segmentation (example in Fig. 21) are detected as false negatives, objects missing in the segmentation are detected as false negatives and there are also a few false negatives where the objects are not connected in the segmentation, although they should be. Errors caused by the algorithm itself and not the input data are mostly due to V-shaped branched objects (the red objects in Fig. 14(d)).

We also evaluated the sensitivity of shape classification to different IOU thresholds and found that overall it is not very sensitive to the choice of the threshold. Performance slightly decreases when the IOU threshold increases (up to 0.75), due to small components, which are not taken into account when the threshold is lower than approx. 0.75. The performance is stable up to approx. 0.9, when it drops more significantly because of segmentation errors - some noisy components are not counted as TP anymore.

Both presented algorithms are nevertheless extracting most of the objects of interest; as they are general, they could also be applied also to different domains with tubular objects such as vesicles, neurons, etc.

6. Conclusion

For a comprehensive analysis of volumetric microscopy data, we need solutions that combine multiple steps: robust segmentation, quality reconstruction, and quantification of important parameters which in addition to simple quantities such as count, density, and position of the observed compartments, also includes more complex ones related to the analysis of component shapes. In the paper, we addressed several of these challenges and next to the proposed methods, we are also releasing a public dataset to encourage further research in the field.

We presented methods for segmentation and reconstruction of FVs which is to our knowledge the first of its kind. We presented a method for segmentation of the GA, which relies on approximately labeled ground truth, thus decreasing the time needed for creating manual annotations; a similar methodology can be applied to many related segmentation problems. We presented a method for reconstruction of mitochondria that improves on the original segmentation by separating the (incorrectly) merged components. Such object-wise reconstruction is necessary, as it allows for reliable quantification of objects, such as counting, measuring volume, etc. We also presented two methods and ground truth data for analyzing the shape of mitochondria and labeling them as contacting or branched. Such properties are important to analyze as they can indicate different physiologically significant states such as stress or spinal injuries. To our knowledge, such methods have not yet been proposed. We also showed that the presented methods can generalize well to data outside our test and training sets by showing the reconstruction of a large FIB-SEM volume.

We released a public dataset, which is to our knowledge the first volumetric EM dataset with manual annotation of FVs. Next to the recently published dataset by Wei et al. [38], which provides labels for instance segmentation of mitochondria, our dataset is the only one providing instance labels for several intracellular compartments. A step from semantic segmentation to more complex instance segmentation is crucial to ensure that we can trust in quantification measures derived from reconstructed volumes - in current related works, these measures are usually reported and used for the biomedical analysis, but their quality is not analyzed.

There is still a lot of room for improvements of the presented methods. Our first plans are to improve the reconstruction of FVs, which we will address with proposal-free instance segmentation methods. Inclusion of more diverse datasets will enable training of better segmentation models, which will also positively affect the quality of reconstruction. Since many of the presented methods address challenges not tackled before, we hope that with the paper we set a benchmark that will be improved in the future by us and other researchers.

We will also continue our cooperation with biomedical experts and hope that the work presents a step towards a comprehensive analysis of volumetric microscopy data. By focusing on the urothelial tissue we want to support researches of the connections between quantitative parameters and bladder diseases, aging, and spinal cord injuries for which it has been shown that they are reflected in the altered structures of intracellular compartments.

Declaration of Competing Interest

Authors declare that they have no conflict of interest.

CRediT authorship contribution statement

Manca Žerovnik Mekuč: Methodology, Software, Writing – original draft. **Ciril Bohak:** Writing – review & editing. **Eva Boneš:** Methodology, Software. **Samo Hudoklin:** Data curation, Writing – original draft, Writing – review & editing. **Rok Romih:** Data curation, Writing – review & editing. **Matija Marolt:** Methodology, Software, Writing – original draft, Writing – review & editing.

Acknowledgements

The authors acknowledge the financial support from the Slovenian Research Agency (research core funding no. P3-0108) and the support of the NVIDIA Corporation with donation of a Titan V GPU used for this research. We thank Bruno M. Humbel and Caroline Kizilyaprak for their dual-beam expertise. We acknowledge the work of Žiga Lesar, who is the author of the VPT visualization tool that we used to visualize the entire volume.

References

- [1] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, P. Walter, *Molecular Biology of the Cell: Sixth International Student Edition*, 500 Tips, Garland Science, Taylor and Francis Group, 2014. <https://books.google.com/books?id=1ZUDoQEACAAJ>.
- [2] S.J. Annesley, P.R. Fisher, Mitochondria in health and disease, *Cells* 8 (7) (2019), doi:10.3390/cells8070680. <https://www.mdpi.com/2073-4409/8/7/680>.
- [3] S. Hudoklin, D. Zupancic, R. Romih, Maturation of the golgi apparatus in urothelial cells, *Cell Tissue Res.* 336 (3) (2009), doi:10.1007/s00441-009-0779-9.
- [4] J. Li, E. Ahat, Y. Wang, Golgi structure and function in health, stress, and diseases, in: M. Kloc (Ed.), *The Golgi Apparatus and Centriole: Functions, Interactions and Role in Disease*, Springer International Publishing, Cham, 2019, pp. 441–485, doi:10.1007/978-3-030-23173-6_19.
- [5] F.M. Platt, B. Boland, A.C. van der Spoel, *Lysosomal storage disorders: the cellular impact of lysosomal dysfunction*, *J. Cell Biol.* 199 (5) (2012) 723–734.
- [6] M.G. Dalghi, N. Montalbetti, M.D. Carattino, G. Apodaca, The urothelium: life in a liquid environment, *Physiol. Rev.* 100 (4) (2020) 1621–1705, doi:10.1152/physrev.00041.2019.
- [7] C.-C.A. Hu, F.-X. Liang, G. Zhou, L. Tu, C.-H.A. Tang, J. Zhou, G. Kreibich, T.-T. Sun, Assembly of urothelial plaques: tetraspanin function in membrane protein trafficking, *Mol. Biol. Cell* 16 (9) (2005) 3937–3950, doi:10.1091/mbc.e05-02-0136.
- [8] S. Hudoklin, K. Jezernik, J. Neumüller, M. Pavelka, R. Romih, Urothelial plaque formation in post-golgi compartments, *PLoS One* 6 (8) (2011) 1–9, doi:10.1371/journal.pone.0023636.
- [9] S. Hudoklin, K. Jezernik, J. Neumüller, M. Pavelka, R. Romih, Electron tomography of fusiform vesicles and their organization in urothelial cells, *PLoS One* 7 (3) (2012) 1–8, doi:10.1371/journal.pone.0032935.
- [10] M.E. Kreft, S. Hudoklin, K. Jezernik, et al., Formation and maintenance of blood–urine barrier in urothelium, *Protoplasma* 246 (1–4) (2010) 3–14.
- [11] S.T. Trusckel, D.R. Clayton, J.M. Beckel, J.G. Yabes, Y. Yao, A. Wolf-Johnston, L.A. Birder, G. Apodaca, Age-related endolysosome dysfunction in the rat urothelium, *PLoS One* 13 (6) (2018), doi:10.1371/journal.pone.0198817.
- [12] X.-R. Wu, X.-P. Kong, A. Pellicer, G. Kreibich, T.-T. Sun, Uroplakins in urothelial biology, function, and disease, *Kidney Int.* 75 (11) (2009) 1153–1165.
- [13] M. Cantoni, L. Holzer, Advances in 3D focused ion beam tomography, *MRS Bull.* 39 (4) (2014) 354–360.
- [14] B. Titze, C. Genoud, Volume scanning electron microscopy for imaging biological ultrastructure, *Biol. Cell* 108 (11) (2016) 307–323.
- [15] M.Ž. Mekuč, C. Bohak, S. Hudoklin, B.H. Kim, M.Y. Kim, M. Marolt, et al., Automatic segmentation of mitochondria and endolysosomes in volumetric electron microscopy data, *Comput. Biol. Med.* 119 (2020) 103693.
- [16] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: learning dense volumetric segmentation from sparse annotation, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2016, pp. 424–432, doi:10.1007/978-3-319-46723-8_49.
- [17] F. Milletari, N. Navab, S.-A. Ahmadi, V-Net: fully convolutional neural networks for volumetric medical image segmentation, in: *2016 Fourth International Conference on 3D Vision (3DV)*, IEEE, 2016, pp. 565–571, doi:10.1109/3DV.2016.79.
- [18] K. Kamnitsas, C. Ledig, V.F. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, *Med. Image Anal.* 36 (2017) 61–78, doi:10.1016/j.media.2016.10.004.
- [19] W. Li, G. Wang, L. Fidon, S. Ourselin, M.J. Cardoso, T. Vercauteren, On the compactness, efficiency, and representation of 3D convolutional networks: brain parcellation as a pretext task, in: *International Conference on Information Processing in Medical Imaging*, Springer, 2017, pp. 348–360, doi:10.1007/978-3-319-59050-9_28.

- [20] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. van der Laak, B. van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (December 2012) (2017) 60–88, doi:10.1016/j.media.2017.07.005.
- [21] F. Isensee, P.F. Jaeger, S.A. Kohl, J. Petersen, K.H. Maier-Hein, nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, *Nat. Methods* 18 (2) (2021) 203–211.
- [22] A. Müller, D. Schmidt, C.S. Xu, S. Pang, J.V. D'Costa, S. Kretschmar, C. Münster, T. Kurth, F. Jug, M. Weigert, et al., 3D fib-sem reconstruction of microtubule-organelle interaction in whole primary mouse β cells, *J. Cell Biol.* 220 (2) (2020).
- [23] Heinrich, L., Bennett, D., Ackerman, D., Park, W., Bogovic, J., Eckstein, N., Petruncio, A., Clements, J., Xu, C. S., Funke, J., et al., 2020. Automatic whole cell organelle segmentation in volumetric electron microscopy. *bioRxiv*.
- [24] M. Berning, K.M. Boergens, M. Helmstaedter, Segem: efficient image analysis for high-resolution connectomics, *Neuron* 87 (6) (2015) 1193–1206.
- [25] N. Kasthuri, K.J. Hayworth, D.R. Berger, R.L. Schalek, J.A. Conchello, S. Knowles-Barley, D. Lee, A. Vázquez-Reina, V. Kaynig, T.R. Jones, et al., Saturated reconstruction of a volume of neocortex, *Cell* 162 (3) (2015) 648–661.
- [26] V. Kaynig, A. Vázquez-Reina, S. Knowles-Barley, M. Roberts, T.R. Jones, N. Kasthuri, E. Miller, J. Lichtman, H. Pfister, Large-scale automatic reconstruction of neuronal processes from electron microscopy images, *Med. Image Anal.* 22 (1) (2015) 77–88.
- [27] A. Motta, M. Berning, K.M. Boergens, B. Staffler, M. Beining, S. Loomba, P. Hennig, H. Wissler, M. Helmstaedter, Dense connectomic reconstruction in layer 4 of the somatosensory cortex, *Science* 366 (6469) (2019) 1–13.
- [28] A. Vázquez-Reina, M. Gelbart, D. Huang, J. Lichtman, E. Miller, H. Pfister, Segmentation fusion for connectomics, in: 2011 International Conference on Computer Vision, IEEE, 2011, pp. 177–184.
- [29] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980–2988, doi:10.1109/ICCV.2017.322.
- [30] M. Dong, D. Liu, Z. Xiong, X. Chen, Y. Zhang, Z.-J. Zha, G. Bi, F. Wu, Instance segmentation from volumetric biomedical images without voxel-wise labeling, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 83–91.
- [31] Kopelowitz, E., Engelhard, G., 2019. Lung nodules detection and segmentation using 3Dmask-rcnn. *arXiv preprint arXiv:1907.07676*
- [32] Y. Lei, Z. Tian, S. Kahn, W.J. Curran, T. Liu, X. Yang, Automatic detection of brain metastases using 3D mask R-CNN for stereotactic radiosurgery, in: *Medical Imaging 2020: Computer-Aided Diagnosis*, vol. 11314, International Society for Optics and Photonics, 2020, p. 113142X.
- [33] R. Bermudez-Chacon, P. Marquez-Neila, M. Salzmann, P. Fua, A domain-adaptive two-stream U-Net for electron microscopy image segmentation, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018, pp. 400–404, doi:10.1109/ISBI.2018.8363602.
- [34] M.G. Haberl, C. Churas, L. Tindall, D. Boassa, S. Phan, E.A. Bushong, M. Madany, R. Akay, T.J. Deerinck, S.T. Peltier, M.H. Ellisman, CDeep3M-plug-and-play cloud-based deep learning for image segmentation, *Nat. Methods* 15 (9) (2018) 677–680, doi:10.1038/s41592-018-0106-z.
- [35] J. Liu, L. Li, Y. Yang, B. Hong, X. Chen, Q. Xie, H. Han, Automatic reconstruction of mitochondria and endoplasmic reticulum in electron microscopy volumes by deep learning, *Front. Neurosci.* 14 (2020) 599–611.
- [36] I. Oztel, G. Yolcu, I. Ersoy, T. White, F. Bunyak, Mitochondria segmentation in electron microscopy volumes using deep convolutional neural network, in: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2017, pp. 1195–1200, doi:10.1109/BIBM.2017.8217827.
- [37] J. Roels, J. Hennies, Y. Saey, W. Philips, A. Kreshuk, Domain adaptive segmentation in volume electron microscopy imaging, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp. 1519–1522, doi:10.1109/ISBI.2019.8759383.
- [38] D. Wei, Z. Lin, D. Barranco, N. Wendt, X. Liu, W. Yin, X. Huang, A. Gupta, W. Jang, X. Wang, I. Arganda-Carreras, J. Lichtman, H. Pfister, Mitoem dataset: large-scale 3D mitochondria instance segmentation from EM images, in: *International Conference on Medical Image Computing and Computer Assisted Intervention*, 2020, pp. 66–76.
- [39] L.A. Pon, E.A. Schon, *Mitochondria*, Elsevier, 2011.
- [40] J. Bereiter-Hahn, M. Vöth, Dynamics of mitochondria in living cells: shape changes, dislocations, fusion, and fission of mitochondria, *Microsc. Res. Tech.* 27 (3) (1994) 198–219, doi:10.1002/jemt.1070270303.
- [41] E.P. Bulthuis, M.J. Adjobo-Hermans, P.H. Willems, W.J. Koopman, Mitochondrial morphofunction in mammalian cells, *Antioxid. Redox Signal.* 30 (18) (2019) 2066–2109, doi:10.1089/ars.2018.7534.
- [42] L. Scorrano, Keeping mitochondria in shape: a matter of life and death, *Eur. J. Clin. Invest.* 43 (8) (2013) 886–893, doi:10.1111/eci.12135.
- [43] J. Liu, W. Li, C. Xiao, B. Hong, Q. Xie, H. Han, Automatic Detection and Segmentation of Mitochondria from SEM Images using Deep Neural Network, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, pp. 628–631, doi:10.1109/EMBC.2018.8512393.
- [44] Casser, V., Kang, K., Pfister, H., Haehn, D., 2018. Fast Mitochondria Segmentation for Connectomics. *arXiv:1812.06024*
- [45] Guay, M. D., Emam, Z. A., Anderson, A. B., Aronova, M. A., Leapman, R. D., 2020. Dense cellular segmentation for EM using 2D-3D neural network ensembles. *BioRxiv*.
- [46] W. Li, J. Liu, C. Xiao, H. Deng, Q. Xie, H. Han, A fast forward 3D connection algorithm for mitochondria and synapse segmentations from serial EM images, *BioData Min.* 11 (1) (2018) 24, doi:10.1186/s13040-018-0183-7.
- [47] A.J. Perez, M. Seyedhosseini, T.J. Deerinck, E.A. Bushong, S. Panda, T. Tasdizen, M.H. Ellisman, A workflow for the automatic segmentation of organelles in electron microscopy image stacks, *Front. Neuroanat.* 8 (November) (2014) 1–13, doi:10.3389/fnana.2014.00126.
- [48] Conrad, R., Narayan, K., 2020. Cem500k—a large-scale heterogeneous unlabeled cellular electron microscopy image dataset for deep learning. *bioRxiv*.
- [49] R. Kikinis, S.D. Pieper, K.G. Vosburgh, 3D slicer: a platform for subject-specific image analysis, visualization, and clinical support, in: F.A. Jolesz (Ed.), *Intraoperative Imaging and Image-Guided Therapy*, Springer New York, 2014, pp. 277–289, doi:10.1007/978-1-4614-7657-3_19.
- [50] T.F. Chan, L.A. Vese, Active contours without edges, *IEEE Trans. Image Process.* 10 (2) (2001) 266–277.
- [51] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, *Int. J. Comput. Vis.* 1 (4) (1988) 321–331.
- [52] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, *Int. J. Comput. Vis.* 22 (1) (1997) 61–79.
- [53] M.A. Abramson, *Pattern Search Algorithms for Mixed Variable General Constrained Optimization Problems*, Rice University, 2002 Ph.D. thesis.
- [54] P. Kollmannsberger, M. Kerschnitzki, F. Repp, W. Wagermaier, R. Weinkamer, P. Fratzl, The small world of osteocytes: connectomics of the lacuno-canalicular network in bone, *New J. Phys.* 19 (7) (2017) 073019.
- [55] M. Kerschnitzki, P. Kollmannsberger, M. Burghammer, G.N. Duda, R. Weinkamer, W. Wagermaier, P. Fratzl, Architecture of the osteocyte network correlates with bone material quality, *J. Bone Miner. Res.* 28 (8) (2013) 1837–1845.
- [56] E. Gibson, W. Li, C. Sudre, L. Fidon, D.I. Shaker, G. Wang, Z. Eaton-Rosen, R. Gray, T. Doel, Y. Hu, et al., Niftnet: a deep-learning platform for medical imaging, *Comput. Methods Programs Biomed.* 158 (2018) 113–122, doi:10.1016/j.cmpb.2018.01.025.
- [57] A. Kirillov, K. He, R. Girshick, C. Rother, P. Dollar, Panoptic segmentation, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.