



Automatic segmentation of mitochondria and endolysosomes in volumetric electron microscopy data

Manca Žerovnik Mekuč^{a,*}, Ciril Bohak^a, Samo Hudoklin^b, Byeong Hak Kim^{c,d}, Rok Romih^b, Min Young Kim^c, Matija Marolt^a

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

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

^c School of Electronics Engineering and Research Center for Neurosurgical Robotic System, Kyungpook National University, 41566 Daegu, South Korea

^d Hanwha Systems Corporation, Optronics Team, 1gongdan-ro, 39376 Gumi, South Korea



ARTICLE INFO

Keywords:

Intracellular compartments
Segmentation
Volumetric electron microscopy data
Mitochondria
Endosomes
Lysosomes
Endolysosomes
Urothelium
Deep learning

ABSTRACT

Automatic segmentation of intracellular compartments is a powerful technique, which provides quantitative data about presence, spatial distribution, structure and consequently the function of cells. With the recent development of high throughput volumetric data acquisition techniques in electron microscopy (EM), manual segmentation is becoming a major bottleneck of the process. To aid the cell research, we propose a technique for automatic segmentation of mitochondria and endolysosomes obtained from urinary bladder urothelial cells by the dual beam EM technique. We present a novel publicly available volumetric EM dataset – the first of urothelial cells, evaluate several state-of-the-art segmentation methods on the new dataset and present a novel segmentation pipeline, which is based on supervised deep learning and includes mechanisms that reduce the impact of dependencies in the input data, artefacts and annotation errors. We show that our approach outperforms the compared methods on the proposed dataset.

1. Introduction

Eukaryotic cells are divided into numerous membrane enclosed compartments, or organelles [1]. Mitochondria produce most of the ATP in the cell, but they are also involved in many other cell functions [2]. Endosomes are intracellular compartments of endocytotic pathway that transport material from the plasma membrane to lysosomes [1]. Since degradation of endocytosed material takes place in late endosomes and lysosomes, we use thereafter the combined term ‘endolysosomes’. All these compartments are highly dynamic and plastic, constantly undergoing fusions and fissions, and moving within the cell, which reflect physiologic states and/or differentiation stages of a cell. Since these processes are important for understanding of diseases at the subcellular level [3,4], a robust pipelines for automatic segmentation of intracellular compartments are needed.

In the urothelium, an epithelium covering urinary bladder, cells undergo a unique differentiation from the basal to the superficial cell layer [5]. Normal superficial cells, called umbrella cells, synthesize large amounts of specialized apical plasma membrane that forms a blood–urine permeability barrier [6]. Numerous intracellular compartments, including mitochondria and endolysosomes, contribute to

maintaining the barrier [5,7–9]. Various diseases of the urinary bladder compromise the barrier [10]. Our studies have shown that bladder cancers or various types of cystitis alter the synthesis, transport and degradation of the apical plasma membrane [11–13], but changes in the intracellular compartments in large cell volumes have not yet been studied.

Most intracellular compartments are below or at the resolution limit of light microscopes, therefore their ultrastructural studies are only feasible by electron microscopy (EM). To study three-dimensional (3D) ultrastructure of intracellular compartments and their spatial and temporal distribution at nanometre resolution, two EM techniques are particularly suitable [14]. Electron tomography, which is performed with the transmission electron microscope, has voxel dimensions 1–10 nm; however the volumes are limited to $\approx 1 \mu\text{m}^3$, which represent only a very limited part of a cell (volume of a single umbrella cell is approximately $50.000 \mu\text{m}^3$) [15,16]. On the other hand, dual beam microscopy, which combines a focused ion beam and scanning electron microscope (FIB-SEM), enables voxel dimensions 10 nm–1 μm , but volumes of material studied are in the range ≈ 1 – $50 \mu\text{m}^3$ [16,17]. The FIB-SEM obtains a stack of serial sections by repeated milling of thin

* Corresponding author.

E-mail addresses: manca.zerovnik-mekuc@fri.uni-lj.si (M. Žerovnik Mekuč), ciril.bohak@fri.uni-lj.si (C. Bohak), samo.hudoklin@mf.uni-lj.si (S. Hudoklin), byeonghak81.kim@hanwha.com (B.H. Kim), rok.romih@mf.uni-lj.si (R. Romih), minykim@knu.ac.kr (M.Y. Kim), matija.marolt@fri.uni-lj.si (M. Marolt).

layers of material by a focused ion beam and acquiring of micrographs of the exposed inner surfaces. The result is a large set of volumetric data on intracellular compartments, which needs to be segmented in order to understand it in a context of cell function.

The manual segmentation of various intracellular compartments of interest on hundreds or thousands of micrographs is very time consuming and prone to bias. Therefore, research on methods for automatic segmentation of microscopy data has recently flourished. As in other image analysis fields, many recent works are based on deep convolutional neural networks (CNNs), which are in most of the tasks outperforming the traditional approaches [18]. In 2015 Ronneberger et al. [19] proposed an architecture called the U-Net, designed specifically for two-dimensional (2D) medical image segmentation. The main idea of the U-Net is to incorporate local and larger contextual information from the input image. Based on this concept, many architectures were proposed for volumetric data. Çiçek et al. [20] proposed the 3D U-Net architecture, Milletari et al. [21] proposed the V-Net architecture as an extension to the U-Net layout. At almost the same time, Kamnitsas et al. [22] presented an architecture named DeepMedic, which is also a 3D CNN, but has a dual pathway architecture that processes multiple scale inputs simultaneously. Li et al. [23] presented a volumetric architecture called HighRes3DNet which uses dilated convolutions.

All of the described architectures are achieving state-of-the-art results in different medical domains. As stated in the review paper [18] which revised 380 deep learning papers from the medical image analysis field, the CNNs are currently the top performing approach for many tasks, but the exact architecture is not necessarily the most important determinant in getting a good solution. Authors claim that it is the expert knowledge about the task that can provide advantages, since many researches use the exact same architectures with the same type of networks, but have widely varying results. In the following subsection we describe methods that have been developed specifically for the segmentation of mitochondria and are thus related to our research.

1.1. Mitochondria segmentation

The field of automated segmentation of volumetric EM data has been largely driven by connectomics, the effort to reconstruct neural wiring diagrams, where the CNNs for the segmentation of cellular boundaries were proposed very early [24,25]. Successful approaches have been proposed also for segmentation of synapses which is a similar task to mitochondria segmentation [26,27].

On the basis of related work, several methods have been proposed that are specifically designed for automatic segmentation of mitochondria. Liu et al. [28] presented a method for segmentation from SEM images based on the Mask R-CNN [29]. Their main contribution is in the post-processing of segmentation masks obtained with the deep network. The post-processing is done in three steps: a morphological opening operation is first used to eliminate the small regions and smooth the large regions, a multi-layer (3D) information fusion algorithm is then used to eliminate the mitochondria shorter than a set threshold and finally an algorithm is employed to improve the consistency in the adjacent layers. Combining a deep CNN with post-processing was also proposed by Oztel et al. [30]. They have developed their own CNN architecture where training is done using $32 \times 32 \times 1$ non-overlapping blocks extracted from the training volume in electron microscopy volumes. Blocks are assigned a ground truth label based on the percentage of pixels from mitochondria and non-mitochondria classes. The last fully connected layer of the network outputs two channel mitochondria versus non-mitochondria class scores, that are then converted to binary classification. They also present three steps of post-processing: 2D spurious detection filtering, boundary refinement, and 3D filtering. All of the described approaches are showing promising results but contrary to our method they do not use 3D spatial information in network training.

While all of the described approaches use 2D convolutions, Haberl et al. [31] presented a 3D convolution based approach called the CDeep3M. It is a ready-to-use volumetric segmentation solution employing a cloud-based deep CNN called the DeepEM3D [32]. Results of mitochondria segmentation with DeepEM3D do not outperform state-of-the-art results, however the approach is interesting because it is very robust and achieves good results on different target classes (nuclei, mitochondria, synaptic vesicles, membrane).

Because of small training datasets, a new type of methods based on domain adaption algorithms have arisen. By now, they do not outperform the existing algorithms for mitochondria segmentation, but results are promising. Bermudez-Chacon et al. [33] proposed domain-adaptive two-stream U-Net [33]. This approach use training data from the domain with plenty of training data to improve the segmentation on another domain with less training data. They propose a method with dual U-Net architecture where they use one stream for the source domain and another one for the target domain. The streams are connected so that they share some of the weights. In [34], authors propose the Y-Net architecture which adapts the classical encoder-decoder layout with an added reconstruction decoder in order to align the source and target encoder features. They tested their work transferring knowledge from isotropic FIB-SEM to anisotropic TEM volumes as well as from brain EM images to HeLa cells.

Public datasets for evaluation of mitochondria segmentation are scarce. The most widely adopted datasets were developed by Lucchi [35] and Xiao [36]. Currently, the best approach according to evaluations on Lucchi's dataset is the supervoxel based method of the same authors [37]. They used a nonlinear RBF-SVM classifier to segment mitochondria in 3D and 2D data. It is one of the rare approaches which does not rely on the CNNs. The best approach according to evaluations on the Xiao's dataset is the deep learning approach which exploits 3D spatial information proposed by the same authors [36]. They used a variant of the 3D U-Net with residual blocks. To solve the problem of vanishing gradients during training, they injected auxiliary classifier layers into the hidden layers.

Segmentation of mitochondria was addressed also for fluorescence microscopy data where the target structures were tagged with the use of fluorescence contrasting. Some of the most recent advances are presented in [38], where iterative deep learning workflows allow for generation of initial high-quality three-dimensional segmentations, which are then used as annotations for training deep learning models.

1.2. The aim

The motivation for the work presented in this paper is to further the research on segmentation of intracellular compartments. We propose a method for automatic segmentation of two types of intracellular compartments — mitochondria and endolysosomes. We propose a novel, publicly available urothelial FIB-SEM dataset (the UroCell dataset), which will enlarge the variety of available datasets with annotated intracellular compartments.

The new dataset is to our knowledge the first public dataset for segmentation of mitochondria and endolysosomes which is not obtained from brain tissue, as well as the first isotropic dataset with labelled endolysosomes.

We evaluate several state-of-the-art approaches to medical data segmentation on our dataset, and propose a new CNN-based segmentation pipeline, which achieves state-of-the-art results for mitochondria and endolysosomes segmentation. In our approach, we introduce techniques to increase the robustness of segmentation by reducing the problem of class imbalance, reducing the impact of varying brightness/contrast and image quality in different parts of the dataset and reducing the impact of unreliable annotations. By making the segmentation pipeline more robust, we demonstrate that our approach can also yield state-of-the-art results on other public isotropic dataset as well.

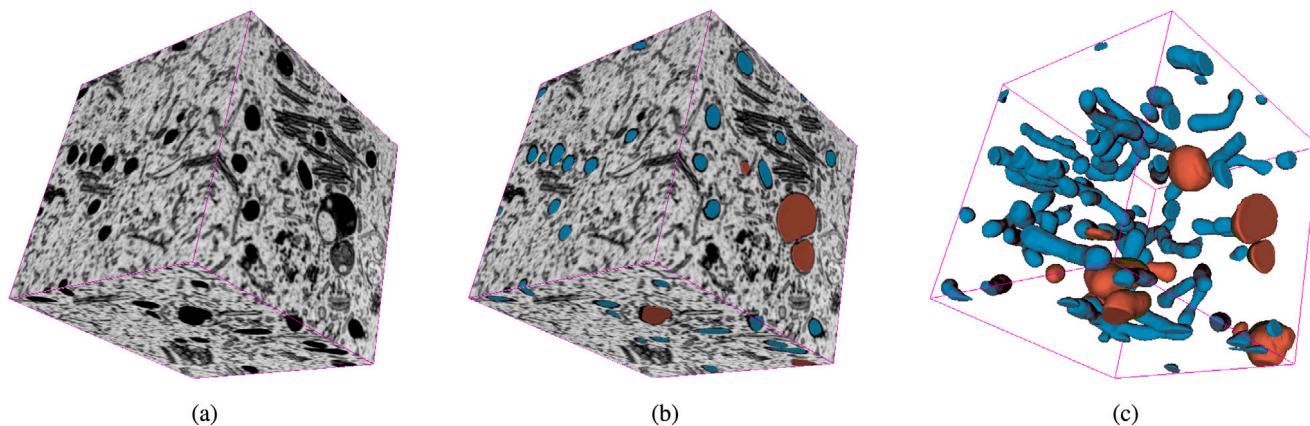


Fig. 1. A three-dimensional cell sub-volume from the UroCell dataset and the corresponding manual annotations. The size of the sub-volume is $256 \times 256 \times 256$ voxels. (a) Raw sub-volume. (b) Manual annotations of intracellular compartments shown on the sub-volume. Mitochondria are shown in blue, endolysosomes in orange. (c) Manual annotations only, shown with default surface smoothing in the Slicer3D Segment Editor [39].

2. Materials and methods

In this section, we describe the novel *UroCell* dataset, and outline our proposed method for segmentation of intracellular compartments.

2.1. The *UroCell* dataset

We obtained tissue samples from urinary bladders of 6–8 weeks old healthy male C57BL/6J mice in accordance with European guidelines and Slovenian legislation. The experiments were approved by the Veterinary Administration of the Slovenian Ministry of Agriculture and Forestry (permit no. U34401-6/2015/4) in compliance with the Animal Health Protection Act and the Instructions for Granting Permits for Animal Experimentation for Scientific Purposes. The volumetric data were obtained from the tissue with a FIB-SEM dual beam electron microscope.

Briefly, urinary bladders were isolated and immediately cryo-fixed with a CPC device (Leica), freeze-substituted in AFS (Leica) with 2% OsO₄ in acetone, and embedded in Epon as described previously [40]. To locate the region of interest, which were terminally differentiated cells of the bladder epithelium (umbrella cells), Epon ultrathin sections were inspected with the Philips CM100 transmission electron microscope and the location of cells was correlated with their location in the epon block. The epon block was then taken into a dual beam Helios NanoLab 650 microscope (FEI) and the umbrella cell was sectioned with the FIB. The dimensions of each voxel were $x = 5.49$ nm, $y = 5.49$ nm, $z = 15.0$ nm. We binned x and y pixels by 3 to obtain an almost isotropic resolution in all three directions which resulted in 1056 serial sections of 1366×1180 pixels. The voxel dimensions in the dataset are thus approximately $x = 16$ nm, $y = 16$ nm, $z = 15$ nm.

We manually annotated intracellular compartments in 5 sub-volumes of size $256 \times 256 \times 256$ voxels. An example of a sub-volume and related annotations is shown in Fig. 1.

To increase variability, the selected annotated sub-volumes come from different parts of the entire volume and are therefore diverse in terms of contrast, brightness, artefacts and content. An example is shown in Fig. 2. We believe that such diversification of annotated sub-volumes is more appropriate to train and evaluate machine learning approaches (training and test datasets contain entirely different sub-volumes); namely in most of the other existing mitochondria datasets (see Section 3), one volume is split into two parts (the training and the test set), which may consequently be quite similar, as the last slice of the training set is a direct neighbour of the first slice of the test set.

We manually labelled two classes of intracellular compartments in all sub-volumes: mitochondria and endolysosomes. All labels were

revised by a cell biologist with expertise in urothelial biology. Currently, in each volume there are from 15 to 85 mitochondria and approximately 10 endolysosomes. For annotation, we used the open source Slicer3D software [39]. With this paper, we make the UroCell dataset publicly available on the following link: <https://github.com/MancaZerovnikMekuc/UroCell> under the CC-BY-NC-SA 4.0 licence.¹

2.2. The segmentation pipeline

To automatically segment mitochondria and endolysosomes, we propose a novel segmentation pipeline based on an existing CNN architecture. As both target classes are quite similar in texture, but also quite unbalanced (there are many more mitochondria than endolysosomes), we propose to use transfer learning to transition from the problem of segmenting any of the two classes from the background to the problem of segmenting each class separately. Thus, we wish to make use of the available information of the larger mitochondria class to improve the segmentation of endolysosomes.

The proposed segmentation pipeline is shown in Fig. 3, each step is described in more details in the following subsections. We first enhance the contrast of the input volumes (Section 2.3). The contrast enhanced volumes are then passed to the upper branch of the pipeline where we train a convolutional neural network model as a binary classifier between all compartments and the background. We then transfer all of the learned weights, except for the last layer, to the lower branch, where we train a three-class CNN classifier (mitochondria, endolysosomes, background), and thus adapt the weights of the new three-class model. We outline the architecture of the CNN in Section 2.5. To make the model more robust, we also perform balanced window selection and data augmentation before training, as described in Section 2.4.

2.3. Contrast enhancement

Although convolutional neural networks are well capable of learning to recognize the target classes in images and volumes that have varying brightness and contrast, it still helps to adjust for the difference before learning due to the relatively small sizes of the FIB-SEM datasets. We thus employ a contrast enhancement algorithm to preprocess the data as a first step in our pipeline. The contrast enhancement algorithm emphasizes the target intracellular compartments and their inner structure, which together with their shape represents the main feature that helps to distinguish mitochondria from the endolysosomes and also from other intracellular compartments.

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

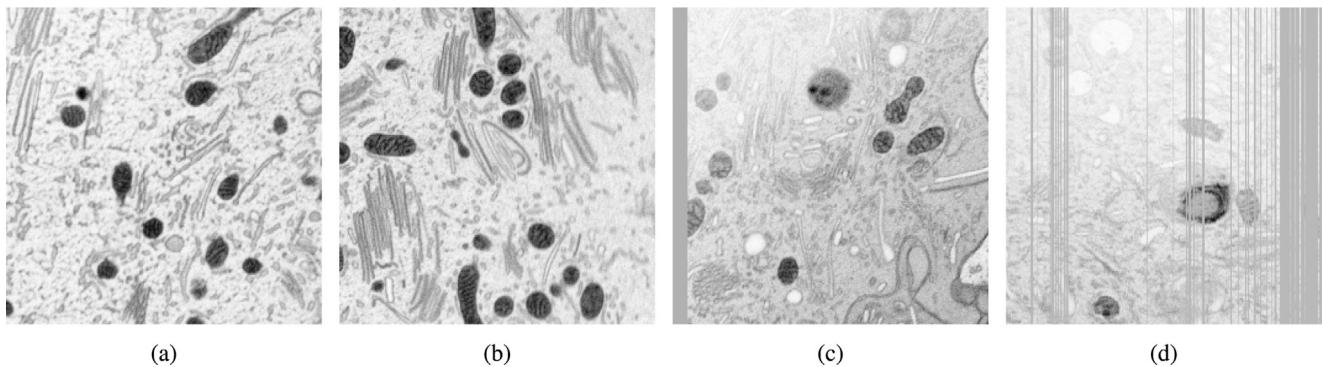


Fig. 2. Slices from two sub-volumes of the UroCell dataset. The slices (a) and (b) are from one sub-volume and the slices (c) and (d) are from the other. Differences in content, brightness, contrast and artefacts are clearly visible.

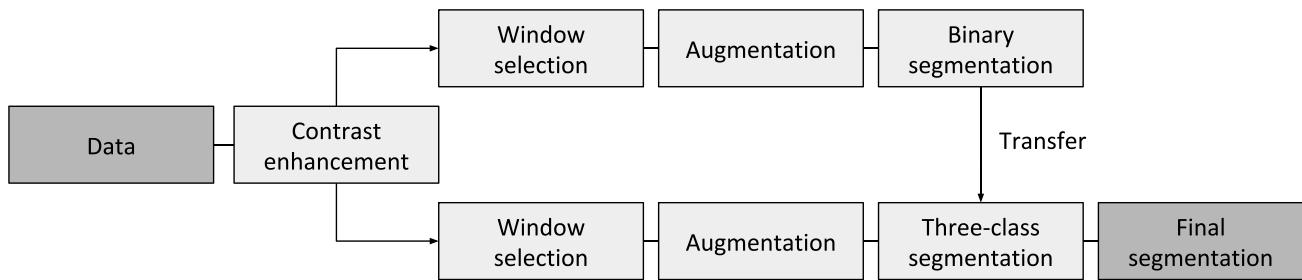


Fig. 3. The proposed pipeline, which consists of two branches and transfer of weights from the upper to lower branch.

Our contrast enhancement approach is based on the adaptive gamma correction with weighting distribution (AGCWD) method [41], which represents an efficient way to enhance the contrast in images. An example of how AGCWD improves contrast is shown in Fig. 4, where 4a and 4d represent original images, and 4b and 4e their enhanced version obtained by the AGCWD method.

Intracellular compartments are already much more clearly visible in the AGCWD enhanced images, however if we inspect the histograms, we can observe that their range is not optimal because of the background and outlier (very bright/dark) voxels. Therefore, we propose an extended AGCWD method that ignores outliers and better rescales the histogram and thus improves the contrast.

The original AGCWD method is as follows. First, the histogram of input intensities is calculated:

$$\text{pdf}(l) = n_l / n_{\text{total}}, \quad (1)$$

where n_l is the number of input values with intensity l and n_{total} the total number of input values. As in the original paper, we can view the histogram as a probability density function of input intensities and thus denote it as $\text{pdf}(l)$.

The weighted probability density function $\text{pdf}_w(l)$ is then defined as:

$$\text{pdf}_w(l) = \text{pdf}_{\max} \left(\frac{\text{pdf}(l) - \text{pdf}_{\min}}{\text{pdf}_{\max} - \text{pdf}_{\min}} \right)^{\alpha}, \quad (2)$$

where α is an adjustable parameter, pdf_{\max} the maximum histogram value and pdf_{\min} the minimum histogram value. The cumulative distribution function, $\text{cdf}_w(l)$ is formulated as:

$$\text{cdf}_w(l) = \frac{\sum_{i=0}^l \text{pdf}_w(i)}{\sum_{l=0}^{\max} \text{pdf}_w(l)}. \quad (3)$$

Finally, the contrast enhanced intensity mapping $T(l)$ is calculated as:

$$T(l) = l_{\max} (l / l_{\max})^{1 - \text{cdf}_w(l)}, \quad (4)$$

where l_{\max} is the maximum intensity of the input and l the intensity of each voxel in the input.

The AGCWD method is susceptible to outliers. Very bright or dark voxels skew the histogram distribution and lead to sub-optimal contrasts, as shown in Figs. 4b and 4e. In order to reduce the effect of outliers, we propose the following simple modification of the algorithm. The modification only considers intensities within the range of 3.5 standard deviations of all intensities. Such range was chosen, because intensities of the target compartments fall well into it, and thus only parts of the volume that contain regions of less interest will be pushed to the edges of the histogram. If we thus consider α and β to be the lower and upper values of intensities that fall within the ± 3.5 standard deviations (we can find them easily from the cumulative distribution function of AGCWD corrected intensities), we extend the final histogram as:

$$T_e(l) = \frac{\text{cdf}(l) - \text{cdf}(\alpha)}{\text{cdf}(\beta) - \text{cdf}(\alpha)} l_{\max}. \quad (5)$$

Values that fall outside of the range of allowable intensities are set to the limits of the range.

The final enhanced slices are shown in Figs. 4c and 4f. They show a much clearer view of the cell structure. We use such enhanced volumes for training, validation and testing of deep learning methods.

2.4. Window selection and data augmentation

At each training iteration, we crop a window corresponding to the target network's receptive field from the input volume. Because of unbalanced distribution of both target classes and the background, we choose the location of the window \mathcal{W} according to the following probabilities:

$$P(\mathcal{W}) = \begin{cases} 0.4 & \text{if } \exists v \in \mathcal{W}, \text{ such that } v \in \mathcal{M} \\ 0.4 & \text{if } \exists v \in \mathcal{W}, \text{ such that } v \in \mathcal{EL} \\ 0.2 & \text{if } \exists v \in \mathcal{W}, \text{ such that } v \in \mathcal{B}, \end{cases} \quad (6)$$

where v is a voxel and $\mathcal{M}, \mathcal{EL}, \mathcal{B}$ sets of voxels labelled as mitochondria, endolysosomes, and background respectively. Such rebalancing boosts

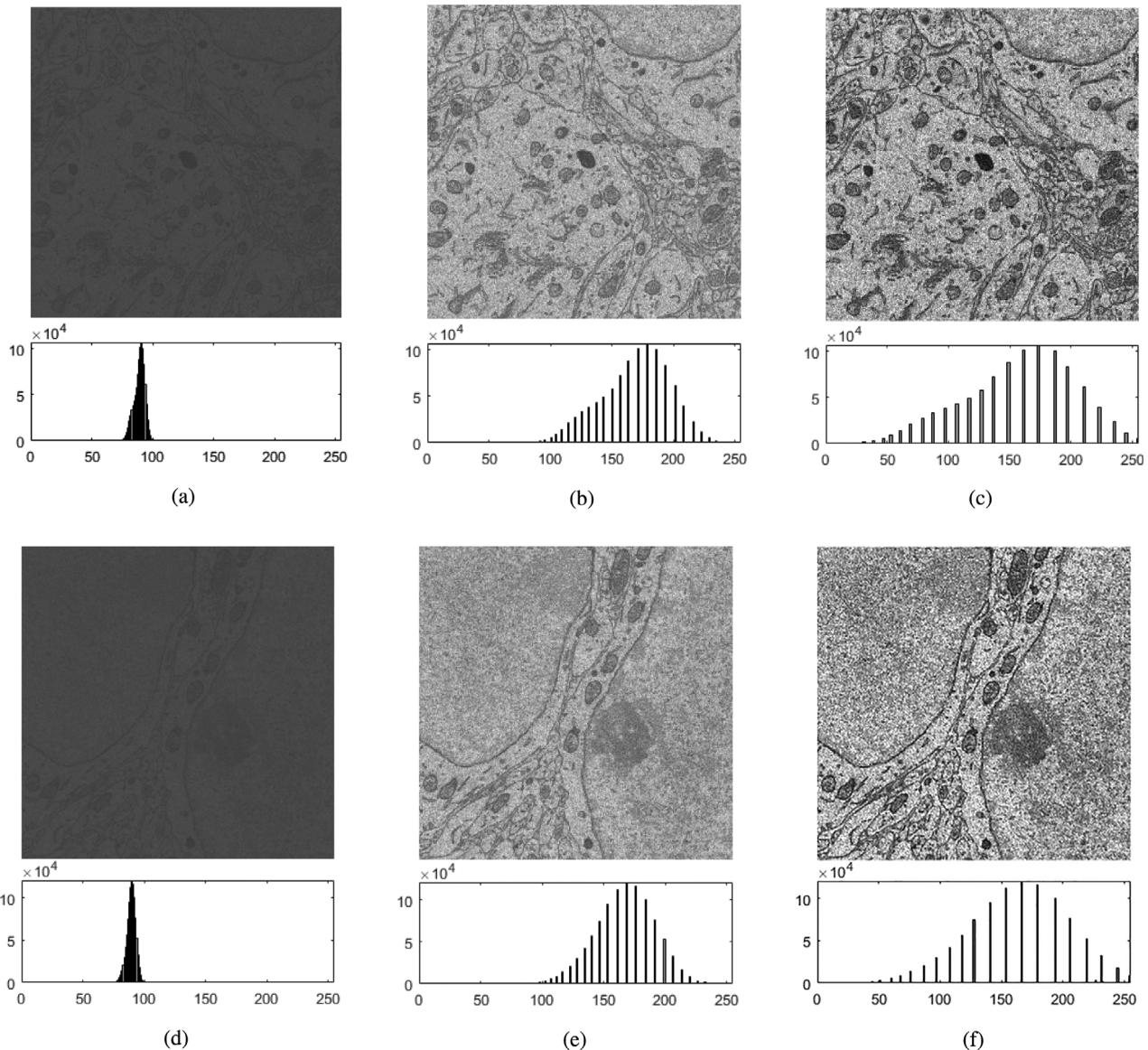


Fig. 4. Examples of contrast enhanced images. (a) and (d) are the original images, (b) and (e) are enhanced using the AGCWD method, (c) and (f) are enhanced using the extended AGCWD method.

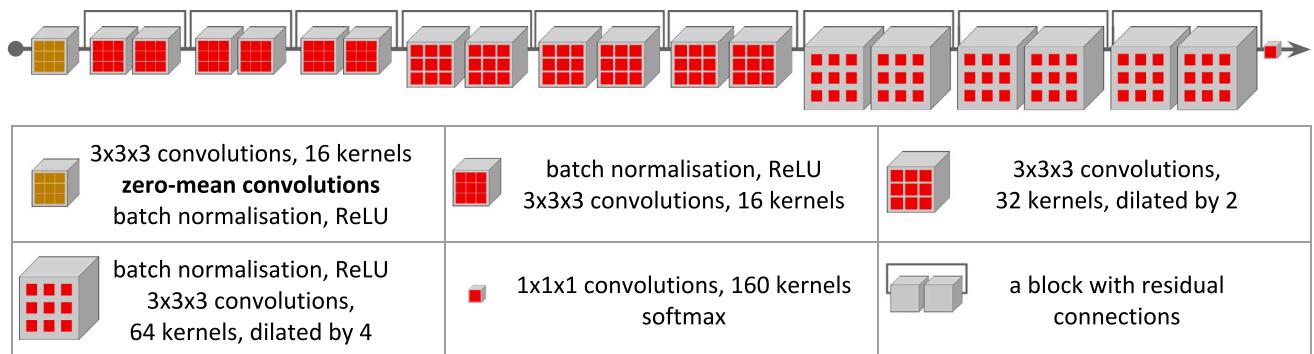


Fig. 5. The HighRes3DZMNet architecture.

both target classes with respect to the background, which is prevalent in the volumes.

As the size of our training dataset is small, increasing the variability of input data is crucial for robust segmentation. To augment the input

data, we perform random flipping of input data over all three dimensions, apply elastic deformations, as used in [21], and random shrinking or expanding of the input data size by up to 10%. All augmentation actions are performed on the fly on 3D windows which we crop from the input volume.

2.5. The proposed network architecture

For segmentation in both branches of the pipeline, we propose a slight modification of the HighRes3DNet [23], which we call HighRes3DZMNet. The authors of the HighRes3DNet architecture initially proposed the architecture for efficient segmentation of volumetric data and they validated it on a brain parcellation task. To incorporate larger contextual and local information, the architecture uses dilated convolutions. The network consists of 20 layers of convolutions. The first seven layers involve $3 \times 3 \times 3$ convolutions, they are followed by six layers where the kernels are dilated by 2, and six layers where the kernels are dilated by 4. At the end there is a gathering convolutional layer with a $1 \times 1 \times 1$ kernel, followed by a softmax decision layer. Each stage contains a batch normalization step. The architecture also uses residual connections between two consecutive convolutional layers except for the first and the last one.

Even though we use the previously described contrast enhancement algorithm to equalize histograms in the sub-volumes, different regions may still exhibit differing levels of brightness, as is also evident in Fig. 2(c) and (d). To diminish the effect that such varying brightness levels may have on segmentation, we modify the input layer of the network by applying zero-mean convolutions, described in the next section. The final proposed architecture is shown in Fig. 5.

2.6. Zero-mean convolutions

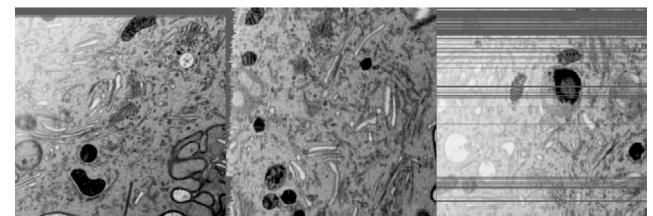
Zero-mean convolutions are a simple and efficient mechanism for making the features of a convolutional neural network layer insensitive to varying levels of brightness in the input. They were initially proposed for processing of sound recordings in order to reduce the impact of varying sound levels [42]. Invariance may be achieved by forcing the coefficients of the learned filters on a given layer to sum to zero:

$$\sum_{i,j,k} K_{i,j,k} = 0 \quad (7)$$

In the upper equation, K is a learned kernel which consists of $i \times j \times k$ coefficients (in our case $3 \times 3 \times 3$). By using kernel constraints, we enforce each of the kernels in the first layer of the CNN, shown in Fig. 5, to have zero mean. In this way, we force the network to learn filters which are based on the differences of voxel values in their receptive fields and not on their levels, which may vary across the volume. Filters are thus also invariant to brightness offsets in the input data, which can be caused by various reasons, such as using a different microscope, different fixation or different settings of the microscope.

2.7. Segmentation masks

During acquisition of FIB-SEM data, different artefacts can appear. We propose to detect and reduce the impact that such artefacts may have on training and evaluation by detecting the artefacts that occur at the borders of the volume. An example is shown in Fig. 6a, which shows the so-called curtaining effect. To detect such artefacts, we simply observe the standard deviation of voxel intensities in each volume dimension, and mask each line of voxels, if its standard deviation is lower than a preset threshold (we use a value of 10^{-2}), as this implies that all voxels in the line have almost the same intensity, indicating an error during acquisition. The detected artefacts are stored in *segmentation masks*, which are volumes of the same size as the input, where the value of a voxel is 0 if it is part of an artefact, or 1 otherwise. An example of the resulting masks is shown in Fig. 6b.



(a)

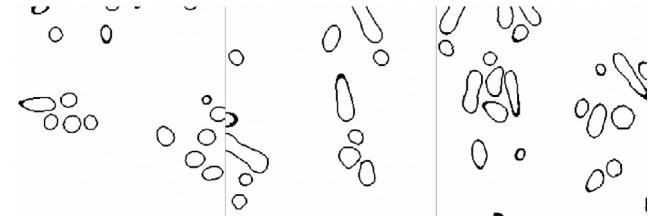


(b)

Fig. 6. (a) X, Y and Z plane intersections of one of the sub-volumes from our dataset. Artefacts are clearly visible. (b) The calculated segmentation masks (black colour represents the masked voxels).



(a)



(b)

Fig. 7. (a) X, Y and Z plane intersections of the annotation of one of the sub-volumes from our dataset. (b) Segmentation masks — object borders are masked as they are difficult to annotate correctly (black colour represents the masked voxels).

As noted in Lucchi et al. [43], human annotations at the borders of volumetric objects may often be inaccurate, as it is often unclear exactly where an object ends. They propose to exclude the borders of the objects in the process of evaluation. We therefore also detect the borders of annotated objects in the manual annotations and add them to the segmentation masks. An example is shown in Fig. 7.

We use the segmentation masks to ignore the masked voxels during evaluation, as their true value is uncertain. We carry this idea further and next to evaluation also modify the training procedure to exclude the masked voxels.

We modify the loss function to include weights from the segmentation masks, which reduce the impact of masked voxels. As we use the Dice loss [21] during training, we modify it as follows:

$$DL = 1 - 2 \frac{\sum_{l=1}^3 \sum_{n=1}^N w_n p_{ln} r_{ln}}{\sum_{l=1}^3 \sum_{n=1}^N w_n (p_{ln} + r_{ln})}, \quad (8)$$

where N is the number of elements in the volume, r_{ln} is the annotated voxel value for l th class, p_{ln} the predicted probability of voxel for the

l th class, and w_n the segmentation mask value of the n th voxel. The segmentation mask thus effectively excludes the masked voxels from the loss.

3. Experiments and results

3.1. Implementation details

We implemented our segmentation pipeline using the NiftyNet framework [47], which is an open source convolutional neural network platform for medical image analysis. For comparison, we implemented several CNN architectures: the U-Net [19], the V-Net [21], DeepMedic [22], HighRes3DNet [23] and our proposed HighRes3DZMNet. To evaluate the approaches on the UroCell dataset, we used 5-fold cross validation, each time using 4 volumes for training and 1 volume for testing, so that each volume from the dataset was used once as the test volume.

To make a fair comparison between the methods, we set most of the hyper-parameters to be the same for all architectures. We employed the L2 regularization with the decay factor set to 0.0001. All of the parameters in the convolutional layers were initialized according to He et al. [48]. The differences among the methods in the training phase were the following: the input window size, the choice of optimizer, the learning rate and the choice of the loss function. We observed that the performance of the networks increases with the size of the input window, therefore for each model we selected a window as large as possible (the limit was GPU memory). We implemented the loss functions as proposed by the authors in their original papers. All the architectures were trained with the Dice loss function.

3.2. Evaluation criteria

For evaluation, we use the following measures, all commonly used in the field of medical image segmentation: sensitivity, specificity and the Dice similarity coefficient (DSC). If we define TP to be the number of true positive voxels (correctly identified target class), FP the number of false positive voxels (target class identified on background), FN the number of false negative voxels (missed target) and TN the number of true negative voxels (correctly classified background), we can define the metrics as follows. The Dice similarity coefficient measures the similarity between annotations and predictions and is defined as:

$$\text{DSC} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}} \quad (9)$$

Sensitivity, also named the true positive rate (TPR) or recall, measures the proportion of correctly classified target voxels with respect to all voxels of the target class:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (10)$$

Specificity, also named the true negative rate (TNR), measures the proportion of correctly classified voxels that do not belong to the target class:

$$\text{TNR} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (11)$$

To compare our approach with results published on other public datasets, we also use the Jaccard index (JI) which is defined as:

$$\text{JI} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (12)$$

With all metrics, a higher value represents a better result.

3.3. Results on the UroCell dataset

We evaluated the proposed method and four other state-of-the-art approaches on the UroCell dataset using the metrics described in the previous section. When calculating the metrics, we also consider the segmentation masks and ignore the masked voxels, as their classes are uncertain. The results are shown in Table 1. The proposed method yields the best results for segmentation of mitochondria, endolysosomes and also the best results if we consider segmentation of both classes.

3.4. Comparable public datasets

Public datasets with annotated intracellular compartments are few. We provide an overview of the available datasets for mitochondria and endolysosomes in Table 2. The proposed UroCell dataset is the only dataset for these two compartments that does not come from a brain region, and next to the dataset of Perez et al. [45] the only one containing annotations for both classes. Apart from the dataset of Lucchi [44], our dataset is also the only isotropic dataset with labels for mitochondria and endolysosomes acquired by the FIB-SEM method.

Since our segmentation model is based on volumetric operations, and does not consider additional preprocessing for anisotropic data, we compare the performance of our segmentation pipeline on the improved version of the dataset initially proposed by Lucchi et al. [37], which contains labels for mitochondria only. The improved version [44] contains a smaller number of annotation errors, which were mostly present because of the initial slice-by-slice manual annotation process, which resulted in very jagged compartment borders in one dimension. We did not change any of the steps or parameters when we applied our method for segmentation of mitochondria on the dataset. The only difference is that since we only classify one class, we did not use transfer learning and only trained and evaluated the upper branch of the proposed pipeline. The results of our pipeline for segmentation of mitochondria are comparable to the results presented in the paper, which proposed the improved version of dataset [44] - the Jaccard index of our approach is **0.90** and is the same as in the mentioned paper.

4. Discussion

Results show that our method yields the best score when segmenting both types of intracellular compartments in the UroCell dataset. If we look at the Dice coefficients presented in Table 1, we can see how different proposed mechanisms and their combinations affect the results. The proposed contrast enhancement scores the highest for the mitochondria class, but has difficulties with endolysosomes as it confuses them often with mitochondria or background. If we compare the mean DSC results of different parts of the pipeline to the baseline HighRes3DNet architecture, we can see that each of them contributes to an improvement and their combination yields the best results.

An example of how the different methods segment both types of compartments in a volume is shown in Figs. 8 and 9. Fig. 8 shows how parts of endolysosomes can be mistaken for mitochondria, while Fig. 9 shows how mitochondria and other compartments can be mistaken for endolysosomes. As the original slices show, the difference in texture of both compartment types is not very large.

Most segmentation errors are made in regions with poor contrast, artefacts or with other cellular compartments, which are misclassified as one of the target classes (e.g. the cellular membrane). As an example, in Fig. 10 we illustrate a slice from the sub-volume with the lowest evaluation score for all methods. The sub-volume has the highest number of artefacts, poor contrast, and also contains parts of the cellular membrane, all of which contribute to segmentation errors. 3D segmentation of the entire sub-volume, where the differences between the methods are clearly visible, is shown in Fig. 11.

As the figures show, with our improved segmentation pipeline we can reduce misclassifications caused by artefacts, poor contrast and the presence of other compartments, however results are still far from perfect for more challenging sub-volumes.

Some of the target objects in very bright areas are not found and there is a number of false positives from parts of the cell which look very alike the target classes. For endolysosomes, a problem also arises when their interior contains large bubbles and during their breakup, when holes appear in the compartments — these are still labelled as the compartments in the manual annotations, while the segmentation methods mostly label them as background, thus yielding false negatives.

Table 1

The performance of different models on the UroCell dataset with an ablation study for proposed mechanisms. The first three columns show the metrics (the Dice similarity coefficient, true positive rate, true negative rate) for the mitochondria (M), the next three columns show the metrics for endolysosomes (EL). The last column shows the mean DSC of both target classes.

Method	DSC-M	TPR-M	TNR-M	DSC-EL	TPR-EL	TNR-EL	Mean DSC
U-Net [19]	0.855	0.825	0.997	0.613	0.631	0.994	0.734
V-Net [21]	0.898	0.873	0.999	0.504	0.572	0.993	0.701
DeepMedic [22]	0.867	0.856	0.997	0.565	0.651	0.997	0.716
HighRes3DNet [23]	0.883	0.857	0.999	0.696	0.743	0.997	0.789
HighRes3DNet - binary classifier for M	0.862	0.843	0.999	—	—	—	—
HighRes3DNet - binary classifier for EL	—	—	—	0.750	0.765	0.998	—
HighRes3DNet merged output of binary classifiers	0.862	0.843	0.999	0.732	0.711	0.998	0.797
HighRes3DNet with transfer learning (TRAN)	0.903	0.881	0.999	0.787	0.750	0.998	0.845
HighRes3DNet with contrast enhancement (CON)	0.954	0.942	0.999	0.660	0.812	0.977	0.807
HighRes3DZMNet	0.882	0.852	0.999	0.727	0.730	0.998	0.804
HighRes3DZMNet with CON	0.937	0.921	0.999	0.801	0.774	0.996	0.869
HighRes3DZMNet with CON and segmentation masks	0.939	0.927	0.999	0.803	0.836	0.995	0.871
Proposed pipeline	0.942	0.921	0.999	0.822	0.852	0.999	0.882

Table 2

The comparable public datasets. For the Lucchi and Kasthuri datasets, we show the improved versions of the original datasets, as proposed in [44].

Dataset	Technique	Region	Voxel size (nm)	Train size	Test size	Labels
Haberl et al. ^a [31]	SBEM	Cerebellum, Habenula	6 × 6 × 40	1024 × 1024 × 80	1024 × 1024 × 20	M
Kasthuri++ ^b [44]	ssEM	Somatosensory Cortex	3 × 3 × 30	1463 × 1613 × 85	1334 × 1553 × 75	M
Lucchi++ ^c [44]	FIB-SEM	Hippocampus	5 × 5 × 5	1024 × 768 × 165	1024 × 768 × 165	M
Perez et al. ^d [45]	SBEM	Hypothalamus	30 × 30 × 8	500 × 500 × 50	500 × 500 × 40	M
Perez et al. ^e [45]	SBEM	Hypothalamus	30 × 30 × 8	500 × 500 × 50	500 × 500 × 40	EL
Li et al. ^f [46]	ATUM-SEM	Cortex	2 × 2 × 50	8416 × 8624 × 31	/	M
The UroCell dataset ^g	FIB-SEM	Urothelium	16 × 16 × 15	(256 × 256 × 256) × 5	Cross-validation	M, EL

^aThe dataset is available on request.

^b<https://casser.io/connectomics/>.

^c<https://casser.io/connectomics/>.

^d<http://www.sci.utah.edu/download/chm/>.

^e<http://www.sci.utah.edu/download/chm/>.

^fCurrently not available (<http://95.163.198.142/MiRA/mitochondria31/>).

^g<https://github.com/MancaZerovnikMekuc/UroCell>.

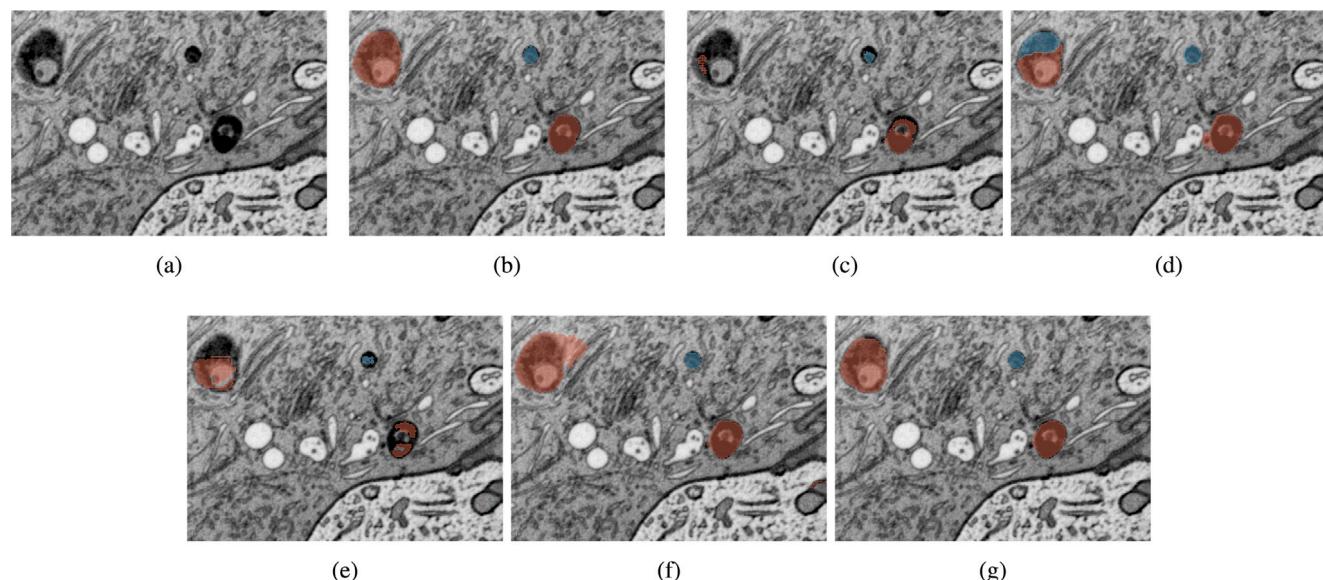


Fig. 8. Segmentation with all of the evaluated methods. The blue colour represents mitochondria, the orange colour represents endolysosomes. Only part of a sub-volume slice is shown. (a) Input data (b) Manual annotations (c) U-Net (d) V-Net (e) DeepMedic (f) HighRes3DNet (g) The proposed method.

Our work shows that further research is needed to improve the robustness of segmentation. We believe that the UroCell dataset that we make public with this paper will contribute to new researches along these lines, as it expands the range of available datasets, as well as introduces more challenges (more compartments, sub-volumes from different parts of the cell with different characteristics). We can see from our evaluation results that the best methods do not achieve the

same accuracy on the UroCell dataset as they do on others, which shows its complexity.

5. Conclusions

With our paper, we make the following contributions. We introduce a novel publicly available dataset with manually labelled intracellular

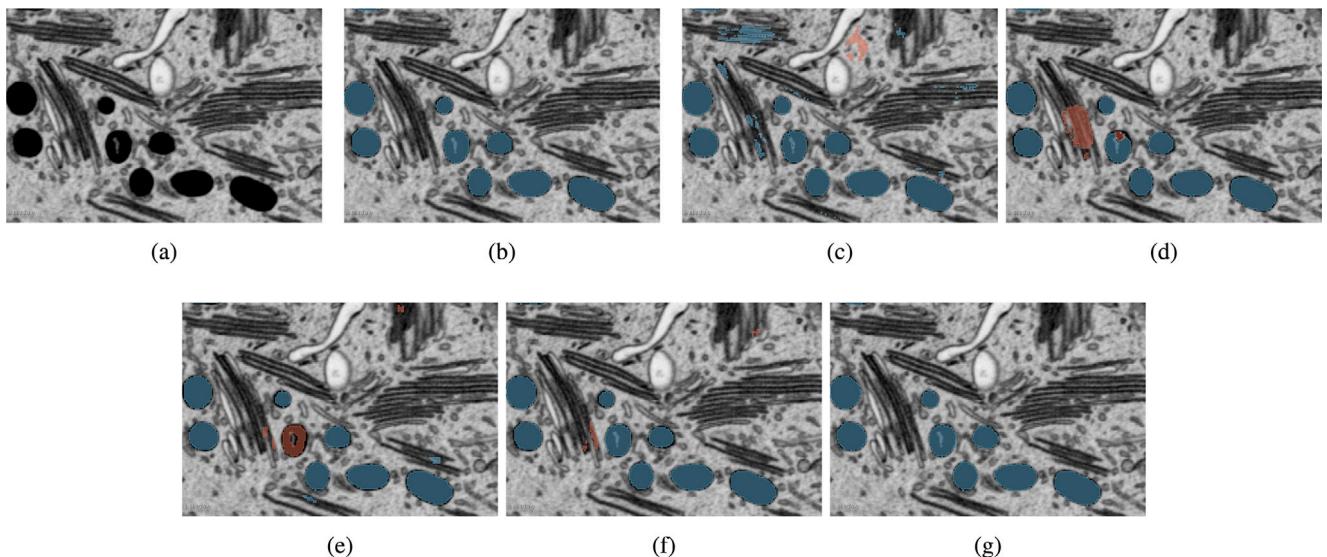


Fig. 9. Segmentation with all of the evaluated methods. The blue colour represents mitochondria, the orange colour represents endolysosomes. Only part of a sub-volume slice is shown. (a) Input data (b) Manual annotations (c) U-Net (d) V-Net (e) DeepMedic (f) HighRes3DNet (g) The proposed method.

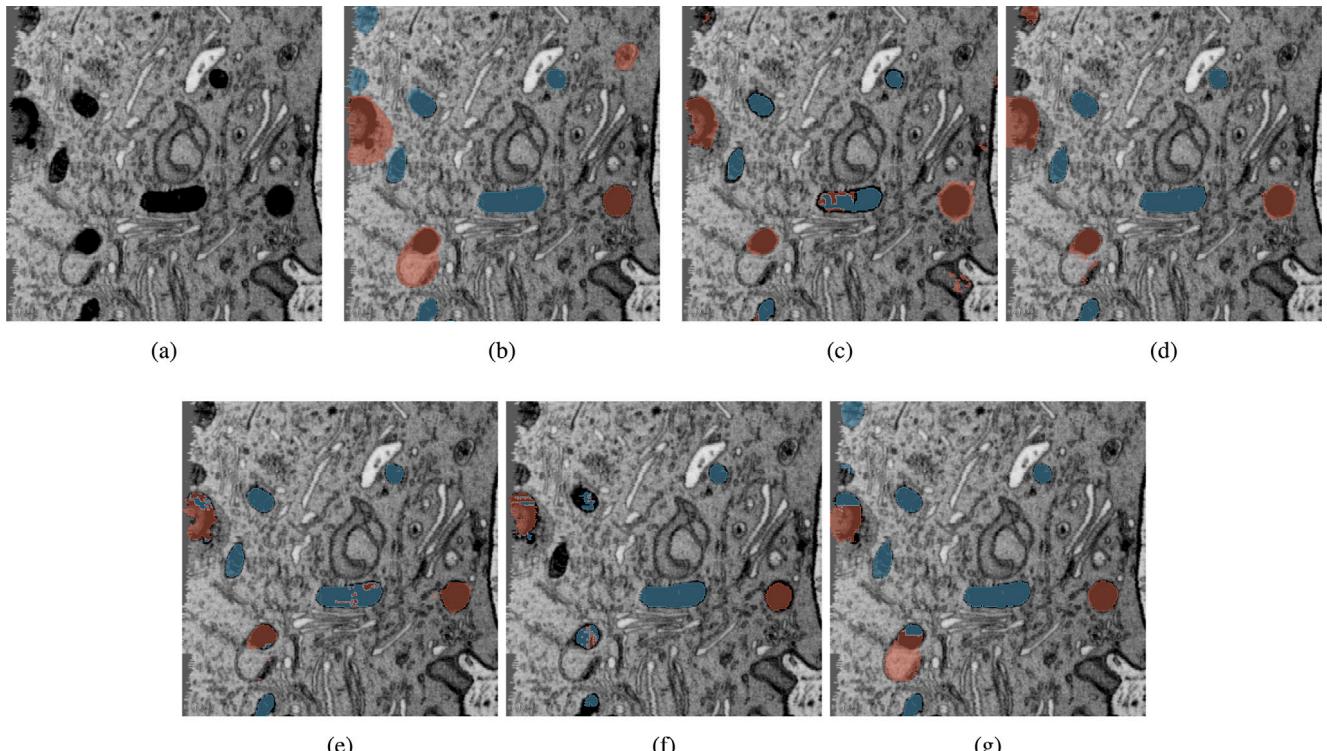


Fig. 10. Segmentation with all of the evaluated methods. The blue colour represents mitochondria, the orange colour represents endolysosomes. (a) Input data (b) Manual annotations (c) U-Net (d) V-Net (e) DeepMedic (f) HighRes3DNet (g) The proposed method.

compartments, which is to our knowledge the first FIB-SEM dataset not obtained from brain tissue which contains labels for both mitochondria and endolysosomes for the same region. The dataset is, in comparison to other public datasets, more diverse, as it consists of five different sub-volumes from different parts of a cell, as well as annotations for two compartment types: mitochondria and endolysosomes.

We evaluated several state-of-the-art deep learning based segmentation methods on the novel dataset, and presented a novel pipeline for segmentation of mitochondria and endolysosomes. With the proposed pipeline, we increase the robustness of segmentation by balanced

sampling, improved contrast and transfer learning. In evaluation, we show that the proposed approach outperforms other state-of-the-art methods.

In our future work, we will be exploring several directions. We will add instance segmentation to our approach, thus isolating individual intracellular compartments. This will enable the analysis of shapes of the isolated compartments, e.g. with spectral shape analysis, which will be useful for improving segmentation accuracy, as well as for studying the characteristics of various compartments. We will also add annotations of new compartment types to the dataset, primarily

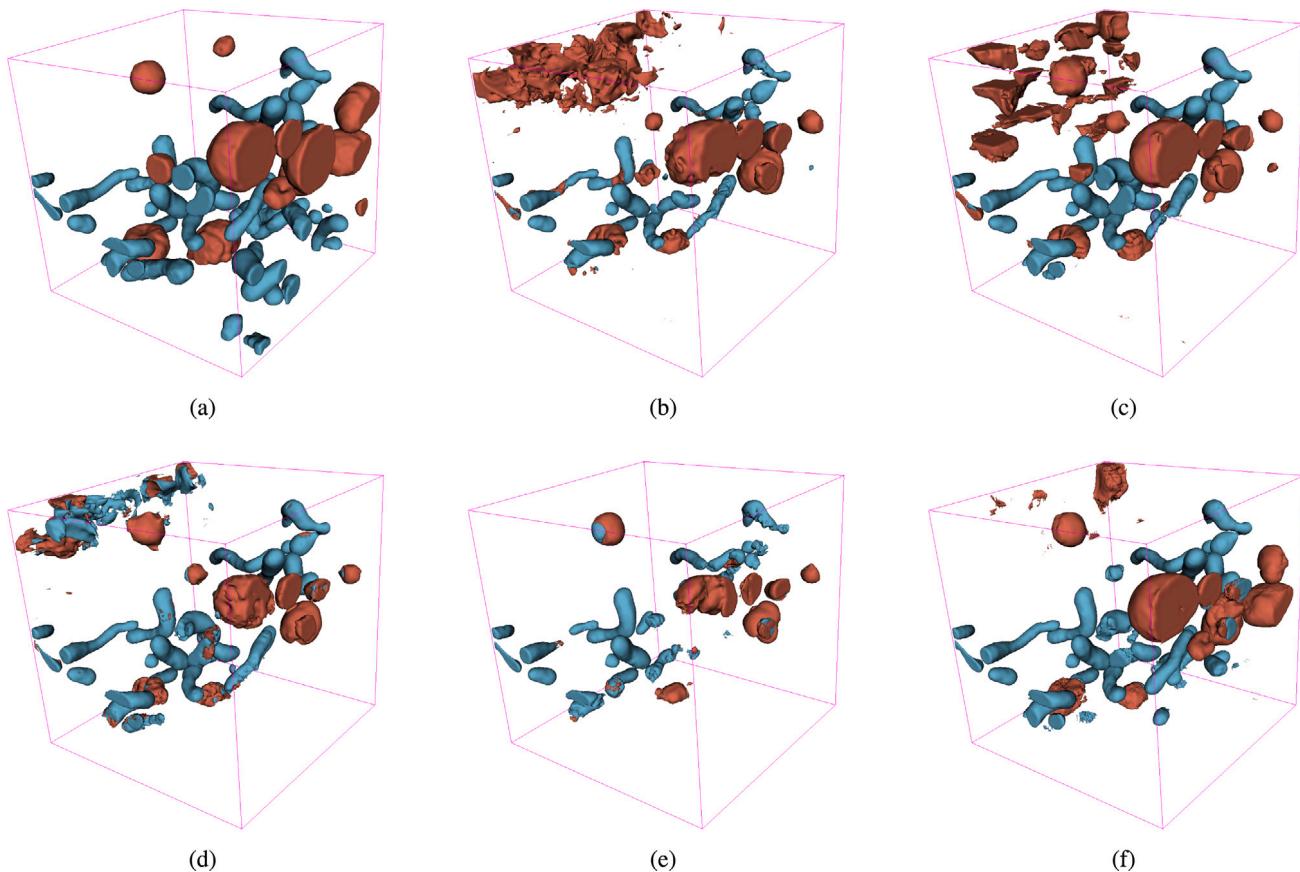


Fig. 11. Segmentation of an entire sub-volume with all of the evaluated methods shown with default surface smoothing in Slicer3D Segment Editor [39]. The blue colour represents mitochondria, the orange colour represents endolysosomes. The chosen sub-volume has the lowest overall evaluation score for all the methods. (a) Manual annotations (b) U-Net (c) V-Net (d) DeepMedic (e) HighRes3DNet (f) The proposed method.

fusiform vesicles and the Golgi apparatus, and develop methods for their segmentation. We hope that with our work we will contribute to a better understanding of cells and their functions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Manca Žerovnik Mekuč: Methodology, Software, Writing - original draft. **Ciril Bohak:** Writing - review & editing. **Samo Hudoklin:** Data curation, Writing - original draft, Writing - review & editing. **Byeong Hak Kim:** Methodology, Software, Writing - original draft. **Rok Romih:** Data curation, Writing - review & editing. **Min Young Kim:** 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 help of all the volunteers who helped us with manual labelling, especially Eva Boneš.

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, in: 500 Tips, Garland Science, Taylor and Francis Group, 2014, URL <https://books.google.com/books?id=1ZUDoQEACAAJ>.
- [2] P. Bhargava, R.G. Schnellmann, Mitochondrial energetics in the kidney, *Nat. Rev. Nephrol.* 13 (10) (2017) 629.
- [3] J. Nunnari, A. Suomalainen, Mitochondria: in sickness and in health, *Cell* 148 (6) (2012) 1145–1159.
- [4] 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.
- [5] 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.
- [6] M.E. Krefl, S. Hudoklin, K. Jezernik, et al., Formation and maintenance of blood–urine barrier in urothelium, *Protoplasma* 246 (1–4) (2010) 3–14.
- [7] P.R. Dodmane, L.L. Arnold, D.E. Muirhead, S. Suzuki, M. Yokohira, K.L. Pennington, B.J. Dave, X. Lu, X.C. Le, S.M. Cohen, Characterization of intracellular inclusions in the urothelium of mice exposed to inorganic arsenic, *Toxicol. Sci.* 137 (1) (2014) 36–46.
- [8] N. Vieira, F.-M. Deng, F.-X. Liang, Y. Liao, J. Chang, G. Zhou, W. Zheng, J.-P. Simon, M. Ding, X.-R. Wu, et al., Snx31: a novel sorting nexin associated with the uroplakin-degrading multivesicular bodies in terminally differentiated urothelial cells, *PLoS One* 9 (6) (2014).
- [9] Y. Liao, D.K. Tham, F.-X. Liang, J. Chang, Y. Wei, P.-R. Sudhir, J. Sall, S.J. Ren, J.U. Chicote, L.L. Arnold, et al., Mitochondrial lipid droplet formation as a detoxification mechanism to sequester and degrade excessive urothelial membranes, *Mol. Biol. Cell* 30 (24) (2019) 2969–2984.
- [10] R. Romih, M. Winder, G. Lee, Recent advances in the biology of the urothelium and applications for urinary bladder dysfunction, *BioMed. Res. Int.* 2014 (2014).
- [11] D. Zupancic, et al., Heterogeneity of uroplakin localization in human normal urothelium, papilloma and papillary carcinoma, *Radiol. Oncol.* 47 (4) (2013) 338–345.
- [12] G. Lee, R. Romih, D. Zupančič, Cystitis: from urothelial cell biology to clinical applications, *BioMed. Res. Int.* 2014 (2014).

- [13] D. Zupančič, M.E. Kreft, et al., Selective binding of lectins to normal and neoplastic urothelium in rat and mouse bladder carcinogenesis models, *Protoplasma* 251 (1) (2014) 49–59.
- [14] K. Miranda, W. Girard-Dias, M. Attias, W. de Souza, I. Ramos, Three dimensional reconstruction by electron microscopy in the life sciences: An introduction for cell and tissue biologists, *Mol. Reprod. Dev.* 82 (7–8) (2015) 530–547.
- [15] R. McIntosh, D. Nicastro, D. Mastronarde, New views of cells in 3D: an introduction to electron tomography, *Trends Cell Biol.* 15 (1) (2005) 43–51.
- [16] M. Cantoni, L. Holzer, Advances in 3D focused ion beam tomography, *MRS Bull.* 39 (4) (2014) 354–360.
- [17] B. Titze, C. Genoud, Volume scanning electron microscopy for imaging biological ultrastructure, *Biol. Cell* 108 (11) (2016) 307–323.
- [18] 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, <http://dx.doi.org/10.1016/j.media.2017.07.005>.
- [19] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241, http://dx.doi.org/10.1007/978-3-319-24574-4_28.
- [20] 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, http://dx.doi.org/10.1007/978-3-319-46723-8_49.
- [21] 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, <http://dx.doi.org/10.1109/3DV.2016.79>.
- [22] 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, <http://dx.doi.org/10.1016/j.media.2016.10.004>.
- [23] 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, http://dx.doi.org/10.1007/978-3-319-59050-9_28.
- [24] S.C. Turaga, J.F. Murray, V. Jain, F. Roth, M. Helmstaedter, K. Briggman, W. Denk, H.S. Seung, Convolutional networks can learn to generate affinity graphs for image segmentation, *Neural Comput.* 22 (2) (2010) 511–538.
- [25] D. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Deep neural networks segment neuronal membranes in electron microscopy images, in: Advances in Neural Information Processing Systems, 2012, pp. 2843–2851.
- [26] L. Heinrich, J. Funke, C. Pape, J. Nunez-Iglesias, S. Saalfeld, Synaptic cleft segmentation in non-isotropic volume electron microscopy of the complete drosophila brain, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 317–325.
- [27] B. Staffler, M. Berning, K.M. Boergens, A. Gour, P. van der Smagt, M. Helmstaedter, Synem, automated synapse detection for connectomics, *Elife* 6 (2017) e26414.
- [28] 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, <http://dx.doi.org/10.1109/EMBC.2018.8512393>.
- [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, <http://dx.doi.org/10.1109/ICCV.2017.322>.
- [30] 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, <http://dx.doi.org/10.1109/BIBM.2017.8217827>.
- [31] 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, <http://dx.doi.org/10.1038/s41592-018-0106-z>.
- [32] T. Zeng, B. Wu, S. Ji, DeepEM3D: approaching human-level performance on 3D anisotropic EM image segmentation, *Bioinformatics* 33 (16) (2017) 2555–2562, <http://dx.doi.org/10.1093/bioinformatics/btx188>.
- [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, <http://dx.doi.org/10.1109/ISBI.2018.8363602>.
- [34] J. Roels, J. Hennies, Y. Saeyns, 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, <http://dx.doi.org/10.1109/ISBI.2019.8759383>.
- [35] A. Lucchi, K. Smith, R. Achanta, G. Knott, P. Fua, Supervoxel-based segmentation of mitochondria in EM image stacks with learned shape features, *IEEE Trans. Med. Imaging* 31 (2) (2012) 474–486, <http://dx.doi.org/10.1109/TMI.2011.2171705>.
- [36] C. Xiao, X. Chen, W. Li, L. Li, L. Wang, Q. Xie, H. Han, Automatic mitochondria segmentation for EM data using a 3D supervised convolutional network, *Front. Neuroanat.* 12 (2018) 92, <http://dx.doi.org/10.3389/fnana.2018.00092>.
- [37] A. Lucchi, P. Marquez-Neila, C. Becker, Y. Li, K. Smith, G. Knott, P. Fua, Learning structured models for segmentation of 2-D and 3-D imagery, *IEEE Trans. Med. Imaging* 34 (5) (2015) 1096–1110, <http://dx.doi.org/10.1109/TMI.2014.2376274>.
- [38] J. Chen, L. Ding, M.P. Viana, M.C. Hendershott, R. Yang, I.A. Mueller, S.M. Rafelski, The allen cell structure segmenter: a new open source toolkit for segmenting 3D intracellular structures in fluorescence microscopy images, 2018, bioRxiv URL <https://www.biorxiv.org/content/early/2018/12/08/491035> <http://dx.doi.org/10.1101/491035>.
- [39] 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, http://dx.doi.org/10.1007/978-1-4614-7657-3_19.
- [40] S. Hudoklin, K. Jezernik, J. Neumüller, M. Pavelka, R. Romih, Urothelial plaque formation in post-golgi compartments, *PLoS One* 6 (8) (2011) 1–9, <http://dx.doi.org/10.1371/journal.pone.0023636>.
- [41] S.-C. Huang, F.-C. Cheng, Y.-S. Chiu, Efficient contrast enhancement using adaptive gamma correction with weighting distribution, *IEEE Trans. Image Process.* 22 (3) (2012) 1032–1041, <http://dx.doi.org/10.1109/TIP.2012.2226047>.
- [42] J. Schlüter, B. Lehner, Zero-mean convolutions for level-invariant singing voice detection, in: Proceedings of the 19th International Society for Music Information Retrieval Conference, 2018, pp. 321–326.
- [43] A. Lucchi, C. Becker, P.M. Neila, P. Fua, Exploiting enclosing membranes and contextual cues for mitochondria segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2014, pp. 65–72, http://dx.doi.org/10.1007/978-3-319-10404-1_9.
- [44] V. Casser, K. Kang, H. Pfister, D. Haehn, Fast mitochondria segmentation for connectomics, 2018, ArXiv.org arXiv:1812.06024 URL <http://arxiv.org/abs/1812.06024>.
- [45] 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, <http://dx.doi.org/10.3389/fnana.2014.00126>.
- [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, <http://dx.doi.org/10.1186/s13040-018-0183-7>.
- [47] E. Gibson, W. Li, C. Sudre, L. Fidon, D.I. Shakir, G. Wang, Z. Eaton-Rosen, R. Gray, T. Doel, Y. Hu, et al., NiftyNet: a deep-learning platform for medical imaging, *Comput. Methods Programs Biomed.* 158 (2018) 113–122, <http://dx.doi.org/10.1016/j.cmpb.2018.01.025>.
- [48] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1026–1034, <http://dx.doi.org/10.1109/ICCV.2015.123>.