Computer Methods and Programs in Biomedicine 219 (2022) 106759

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Deep residual contextual and subpixel convolution network for automated neuronal structure segmentation in micro-connectomics

Chi Xiao[a,b,1], Bei Hong[b,c,1], Jing Liu[b,c], Yuanyan Tang[d], Qiwei Xie[e,**], Hua Han[b,c,f,*]

[a] Key Laboratory of Biomedical Engineering of Hainan Province, School of Biomedical Engineering, Hainan University, China
[b] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China
[c] School of Artificial Intelligence, School of Future Technology, University of Chinese Academy of Sciences, China
[d] Department of Computer and Information Science, University of Macau, China
[e] Data Mining Lab, Beijing University of Technology, China
[f] Chinese Academy of Sciences Center for Excellence in Brain Science and Intelligence Technology, China

ARTICLE INFO

Article history:
Received 15 April 2021
Revised 2 March 2022
Accepted 13 March 2022

Keywords:
Deep learning
Neuronal structure segmentation
Subpixel convolution
Electron microscopy
Micro-Connectomics

ABSTRACT

Background and Objective: The goal of micro-connectomics research is to reconstruct the connectome and elucidate the mechanisms and functions of the nervous system via electron microscopy (EM). Due to the enormous variety of neuronal structures, neuron segmentation is among most difficult tasks in connectome reconstruction, and neuroanatomists desperately need a reliable neuronal structure segmentation method to reduce the burden of manual labeling and validation. Methods: In this article, we proposed an effective deep learning method based on a deep residual contextual and subpixel convolution network to obtain the neuronal structure segmentation in anisotropic EM image stacks. Furthermore, lifted multicut is used for post-processing to optimize the prediction and obtain the reconstruction results. Results: On the ISBI EM segmentation challenge, the proposed method ranks among the top of the leader board and yields a Rand score of 0.98788. On the public data set of mouse piriform cortex, it achieves a Rand score of 0.9562 and 0.9318 in the different testing stacks. The evaluation scores of our method are significantly improved when compared with those of state-of-the-art methods. Conclusions: The proposed automatic method contributes to the development of micro-connectomics, which improves the accuracy of neuronal structure segmentation and provides neuroanatomists with an effective approach to obtain the segmentation and reconstruction of neurons.

1. Introduction

The aim of research in connectomics is to analyze the working mechanism of the brain by studying comprehensive maps of connections within the nervous system. However, the connectome is a comprehensive network composed of tens of billions of neurons and nanoscale synapses. To map the finest connection structures, electron microscopes, with nanometer scale observations, have gradually become an important tool in the study of micro-connectomics [1–4]. In previous studies, Takemura et al. [5] utilized serial section transmission electron microscopy (ssTEM) [6] to sparsely reconstruct the specific neural circuits in the medulla of Drosophila visual lobules, which include 379 neurons and large amounts of chemical synaptic contacts. The reconstruction method is based on semiautomatic neuron recognition and manual verification, which require nearly 15,880 human hours (including 1700 expert hours). In another study, Helmstaedter used SBEM microscopic imaging technology and combined large-scale manual neuron annotation with machine learning-based three-dimensional image segmentation technology to reconstruct the inner network layer of the mouse retina. The reconstruction of 950 neurons and their synaptic connections consumed more than 20,000 human hours, and the reconstruction volume is smaller than 130 μm $\times$ 110 μm $\times$ 80 μm. Without automation, the reconstruction of a 1 mm $\times$ 1 mm $\times$ 1 mm [7] neuronal connectivity map would require more than 1000 times the human effort. To reduce the manual workload, the first

* Corresponding author at: National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China.
** Corresponding author at: Data Mining Lab, Beijing University of Technology, China.
E-mail addresses: qiwei.xie@bjut.edu.cn (Q. Xie), hua.han@ia.ac.cn (H. Han).
[1] Chi Xiao and Bei Hong contributed equally to this work.

https://doi.org/10.1016/j.cmpb.2022.106759

(a) (b) (c)

Fig. 1. Neuronal structure segmentation. (a) EM image; (b) ground truth of the neuronal boundary; (c) corresponding segmentation result.

and fundamental task is to improve the accuracy and efficiency of neuronal structure segmentation and reconstruction methods.

However, this is an extraordinarily challenging task. The challenges illustrated in Fig. 1 mainly lie in the following aspects: (1) the shape and size of neurons in EM images vary greatly, which makes it impossible to obtain the neuronal structure directly. Therefore, the segmentation of neuron boundaries is in demand. (2) Image deformation in the acquisition process may blur the neuronal boundaries between adjacent neurons, which influences the accuracy of recognition. (3) The ultrastructure, such as synapses, mitochondria, vesicles and microtubules, reduce the recognition accuracy of the neuronal boundary and lead to split errors or merge errors. In addition, the accumulated errors in the process of image acquisition and alignment also impose difficulties on the design of an effective neuronal structure segmentation algorithm.

To promote the automation of the segmentation and reconstruction of neuronal structures, Arganda Carra et al. [8] organized the IEEE International Symposium on Biomedical Imaging (ISBI) neuron structure segmentation challenge. Participants can evaluate the performance of segmentation algorithms by submitting their results online [9]. The challenge has attracted the participation of many research teams. Ciresan et al. [10] used a convolutional neural network to train a pixel-level classifier and obtain the probability of each pixel of the neuronal boundary. They then used a filtering model to obtain the final segmentation and achieved the best results in the challenge. Stollenga et al. [11] employed an improved multidimensional LSTM to classify neuron boundaries in EM images, which obtained a better result than that of Ref. [10]. Subsequently, Ronneberger et al. [12] presented a U-shaped full convolutional network (FCN), named U-Net, which utilized skip connections to combine contextual information and further improve the accuracy of neuronal structure segmentation. Quan et al. [13] proposed a full convolutional network based on U-net and a residual network for neuronal structure segmentation and used summation-based connections to construct a deeper network with higher accuracy. Drozdzal et al. [14] presented a pipeline that combined FCN with fully convolutional residual networks (FC-ResNets) to segment EM images. To further utilize multilevel contextual information, Shen et al. [15] proposed a multistage and multirecursive-input FCN for neuronal structure segmentation and suppressed false positive error on intracellular structures.

However, the networks mentioned above suffer from the vanishing gradient problem because of their complex structures, and there are discontinuous boundaries in the output probability maps. Therefore, Beier et al. [16] proposed a neural network named ICv1 to acquire the probability maps of the neuronal boundaries and then used lifted multicut to refine the results, which improved the accuracy of neuronal structure segmentation and ranked near the top of the ISBI EM segmentation challenge leader board. In addition, Vidotto et al. [17] designed a faster training and effective segmentation method that achieved promising results without heavy post-processing. However, all the above methods tend to ignore the influence of downsampling and upsampling, which inevitably reduce the amount of detailed information about the spindly neuronal boundaries. Overall, a state-of-the-art neuronal structure segmentation method requires a deep and effective network to avoid the vanishing gradient problem, multilevel contextual information to differentiate neuronal boundaries and other organelles, and an exceptional post-processing step to refine the segmentation results.

In this paper, our goal is to build an automatic pipeline to obtain the neural structure segmentation and reconstruction in micro-connectomics effectively and accurately. To this end, we first designed an effective method based on a residual contextual and subpixel convolution network (RC-SPCNet) to implement the neuron boundary segmentation. The proposed method applies a spatially efficient residual and multilevel contextual architecture as the fully end-to-end network to improve the prediction of neuronal boundaries and avoid the vanishing gradient problem. In addition, to recover the information of neuronal boundaries as accurately as possible, the subpixel convolution (SPC) approach is applied to upsample the feature maps. To the best of our knowledge, we are the first to utilize the subpixel method to recover the details of neuronal boundaries in EM image stacks, which significantly improves the accuracy of the segmentation. Next, we refined the neuron boundary segmentation and obtain the neuron reconstruction via the lifted multicut post-processing methods. The experimental results on several public data sets suggest that the proposed method surpasses most state-of-the-art methods and approaches the accuracy of human experts. We further applied our method to obtain the neuron structure segmentation and reconstruction in a larger-scale data set to demonstrate the effectiveness of the proposed method in handling large-size data.

2. Methods

In this section, we provide the description and pipeline of the proposed method in detail. As shown in Fig. 2, the automated neuron structure segmentation method for EM data sets consists of neuronal boundary segmentation with the proposed network and neuronal structure segmentation and reconstruction with the post-processing method.

2.1. Overview of key insights

In the design of the proposed method, we aim to achieve four goals: (1) Design a deep and effective network to generate dense predictions and take nonfixed-sized data as input. (2) Incorporate a multilevel receptive field of the feature maps to increase the contextual information and avoid ambiguities in neuronal boundaries and other ultrastructural objects. (3) Utilize an appropriate upsampling operation to recover detailed information about neu-

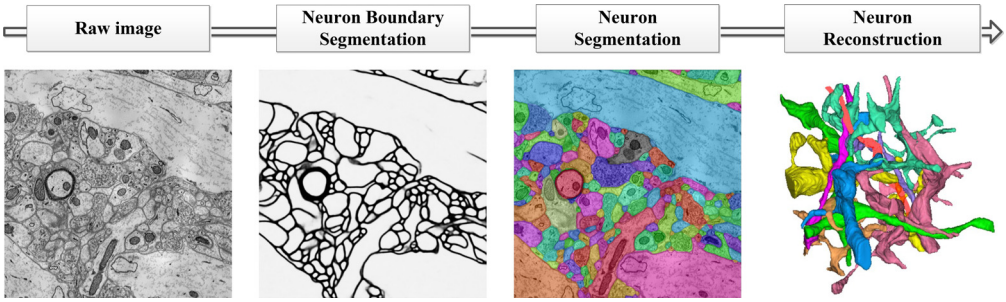

Fig. 2. The pipeline of our method. Left to right: raw images; neuronal boundary segmentation with the proposed network; neuronal structure segmentation and reconstruction with the post-processing method.

ronal boundaries. (4) Adopt an effective post-processing method to further improve the performance and realize the segmentation and reconstruction of neuron structure.

At present, deep neural networks [18–20], particularly residual networks [21], have achieved remarkable results in the study of image classification and segmentation due to the very deep structure and the feature extraction ability. However, training very deep networks is not easy because the gradients vanish exponentially with increasing depth [22]. ResNet38 [23], as an efficient spatial residual network, outperforms very deep ResNet on the ImageNet and PASCAL VOC data sets and is suitable for neuron segmentation in EM images. In our work, we utilized the modified ResNet38 as the fully end-to-end network backbone to generate the prediction of neuronal boundaries.

Since there are large variations in the size of ultrastructures in EM images, the size of the receptive field plays a major role in the task of EM image segmentation. For instance, neuronal boundaries require a small receptive field for classification. However, for the large organelles inside neurons, a larger-scale receptive field is needed to suppress false predictions. Therefore, a network with multiscale receptive fields is necessary. In this case, we designed two skip connections to incorporate different contextual cues and improve the boundary segmentation accuracy.

Considering that the neuronal boundary is a narrow structure with nearly 5 pixels on the width axis, the processes of downsampling and upsampling will inevitably reduce the quantity of available detailed information. In this work, we focused on changing the upsampling model. Compared with traditional upsampling methods, such as deconvolution, subpixel convolution [24,25] fully utilizes all the feature maps and produces dense pixel-wise predictions, which are beneficial for recovering the information of the neuronal boundary and avoiding the loss of detail. To further improve the performance, we used the maximum test-time augmentation (TTA) and state-of-the-art post-processing methods to refine discontinuous boundaries and obtain the segmentation and reconstruction results.

2.2. Network architecture

In this article, we designed an efficient deep neural network to obtain neuronal boundaries. The overview of the RC-SPCNet architecture presented in Fig. 3 mainly consists of two modules: a contracting path with residual-contexture structure and an expansive path with subpixel convolutional layers.

In the contracting path, there are six residual blocks, and each block consists of several residual units. Each residual unit includes two convolutional layers with kernel size 3 × 3 (same padding), followed by batch normalization [26] and exponential linear units [27]. Considering that the boundary of neurons is relatively narrow, an excessive downsampling strategy might destroy the details; therefore, the proposed network only reduces the feature maps at 1/8 resolution of the input size. In Fig. 3, the pooling layer and convolutional layers with stride 2 in residual blocks 1 and 2 are used for downsampling. Moreover, residual blocks 3, 4, 5 and 6 utilize a dilated convolutional layer to increase the receptive field and improve the segmentation performance. Furthermore, we utilized summation-based skip connections [28] to fuse the multiscale information, which helps recognize ambiguities in neuronal boundaries and other ultrastructural objects. The number of skip connections k depends on the number of contextual information channels: when $k = 0$, there are no skip connections, and when $k = 2$, there are two skip connections and the outputs are combined by three-level contextual feature maps.

In the expansive path, we used SPC layers to decode the feature maps. Compared with the deconvolutional layer [29], SPC applies convolution and periodic shuffling operations on the feature maps to obtain a dense pixel-wise prediction map that is efficient, learnable and capable of retaining detailed information. An illustration of the SPC is shown in Fig. 4. Suppose that the width and height of the input image are rW and rH, respectively. The adopted operation first uses convolution to obtain an image I with a shape of $H \times W \times Dr^2$; then, the subpixel ($\mathcal{SP}$) operation rearranges the elements of image I into image $I^{'}$ of $rH \times rW \times D$, where D is the channel of image $I^{'}$ and r is the upsampling factor. The $\mathcal{SP}$ operation can be defined as follows:

$$\mathcal{SP}\{I\}(w,h,d) = I([w/r],[h/r],D\cdot r\cdot \mathrm{mod}(h,r) + D\cdot \mathrm{mod}(w,r) + d), \quad (1)$$

where w, h and d are the width, height and depth indices of the output image $I^{'}$, respectively. $[\cdot]$ represents the floor operation, and mod $(\cdot)$ denotes the modulo operation. In closing, two convolutional layers with dropout and a softmax layer are used to generate pixel-wise predictions.

2.3. Post-processing

The proposed method produces promising segmentation results. However, there are still some discontinuous predictions of neuronal boundaries in ambiguous regions that may lead to merging errors of neurons. We introduce two types of post-processing to refine the segmentation results, namely, the TTA strategy and graph partition algorithm.

To improve the reliability of the boundary probability map and reduce the merge errors among neurons, the TTA strategy is adopted first in the model inference stage. We rotated and flipped each testing image 8 times; the augmented images are then passed into the proposed RC-SPCNet model. Next, the output probability maps are applied via the corresponding reverse transformation and combined. We used an element-wise maximum to combine the output probability maps to ensure maximum boundary continuity and reduce the possibility of merging errors due to boundary fracture.

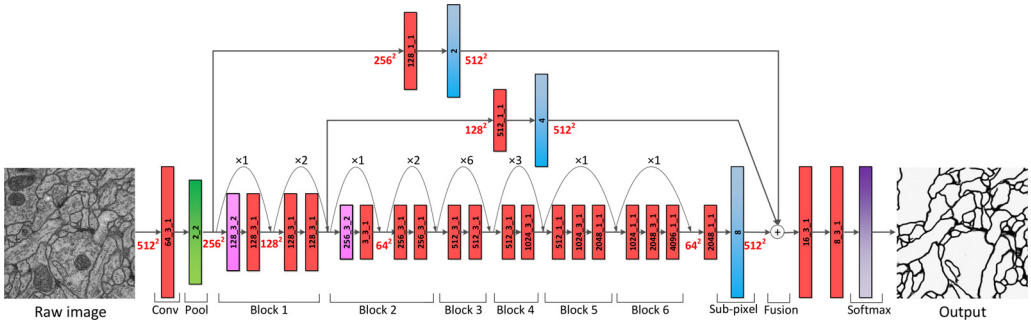

Fig. 3. The proposed architecture of the RC-SPCNet. The red and pink boxes annotated *C_K_S* represent convolutional layers with channels *C*, kernel size *K* and stride *S*. The green box denotes the max-pooling layer, the blue boxes annotated *r* indicate subpixel layers with upsampling factor, and the purple box denotes the softmax layer. The numbers below the straight arrows and curved arrows indicate the size of the feature maps and the repetitions of residual units, respectively.

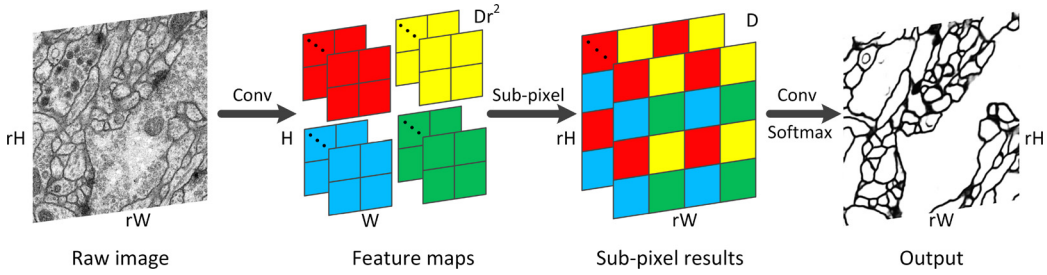

Fig. 4. Illustration of the SPC in neuron segmentation. SPC is in the expansive path of the network and is applied to the decoding process. It utilizes convolution operations to aggregate the feature maps and obtain the dense pixel-wise prediction map.

Subsequently, we utilized lifted multicut to further improve the segmentation performance and obtain the 3D reconstructions. According to Ref. [16], the EM images can be presented as a region adjacency graph, the nodes and edges in this graph denote the regions (super-pixels) and boundaries between the adjacent or non-adjacent regions, respectively. Thus, neuron segmentation problem can be expressed as graph partitioning problem, which can be solved approximately by the lifted multicut method. This task starts from the weighted graphs $G = (V, E)$ and $G' = (V, E')$, where the node V represents super-pixel, regular edges $E \subseteq E'$ link two nodes if and only if they share an image boundary, and lifted edges $E' \setminus E$ denote all pairs of non-neighboring nodes within a predefined distance in the graph. By minimizing the total edge weight, the objective function of the lifted multicut is obtained as follows:

$$\min_{x\in\{0,1\}^{E'}} \sum_{e\in E'} w_e x_e \tag{2}$$

$$\text{subject to} \quad \forall Y \in cylce(G), \forall e \in Y : x_e \leq \sum_{e'\in Y\setminus\{e\}} x_{e'} \tag{3}$$

$$\forall vw \in E' \setminus E, \forall P \in vw{-}paths(G) : x_{vw} \leq \sum_{e\in P} x_e \tag{4}$$

$$\forall vw \in E' \setminus E, \forall C \in vw{-}cuts(G) : 1{-}x_{vw} \leq \sum_{e\in C} (1{-}x_e), \tag{5}$$

where edge label x_e contains binary indicator variables (1 is cut and 0 is uncut) and w_e is the cost of being cut. $Cycles(G)$ represents the set of all cycles in G. Inequality (3) is a constraint to guarantee a consistent solution without "dangling edges". Additional constraints (4) and (5) ensure the lifted edges $vw \in E' \setminus E$ are consistent with the connectivity; for example, $x_{vw} = 0$ if and only if there exists a path (vw-paths) in G connecting v and w.

In this work, the nodes V represent the regions (super-pixels) are generated by the boundary probability map via the distance transform watershed algorithm. The weights w_e encode the attractive strength (positive values) or repulsive strength (negative values) of the regular edges and the lifted edges, which are calculated by the negative log-likelihood formula:

$$w_e = \log \frac{1 - p_e}{p_e}, \tag{6}$$

where p_e represents the similarity of two connected super-pixels belonging to the same neuron, which can be calculated through several features of regular edges and lifted edges [16].

The similarity p_e can be calculated by two steps: feature extraction and classifier training. For the first step, several descriptions of each edge are extracted from the raw images, boundary predictions and corresponding filtered images. To be specific, the feature sets are the same as described in Ref. [16], where the features of regular edges consist of boundary appearance feature, region feature and shape topology feature, and the features of lifted edges consist of connectedness feature, clustering feature and region feature. For the second step, we train two random forests classifiers based on these features, where one learns the similarities probabilities of regular edges, and the other learns the similarities probabilities of lifted edges. In the training process, the edge label of random forests classifiers is generated by the groundtruth of training data, where the edge is labeled as 0 if the two regions connected by this edge belong to the same neuron, otherwise, it is labeled as 1. Next, the features and the corresponding labels are fed into the random forest classifiers to learn whether two superpixels connected by edge represent the same neuron. After the training process, the random forest classifiers are used to infer the p_e in test dataset. Subsequently, the similarity p_e is transformed into edge weights w_e by the negative log-likelihood (6), note that each edge (include regular edge and lifted edge) has p_e corresponds to $e \in E'$ in the optimization model.

Next, the fusion move solver [30] is used to optimize the lifted multicut objective. In particular, the hierarchical solver [31] (a fast greedy search strategy) is utilized to optimize the lifted multicut objective for large volume data, which guarantee that the time

consumption of the proposed method in processing large volume data is within the acceptable range. Finally, nodes are clustered into neurites in 3D by solving the graph partitioning problem. In this case, the post-processing method is capable of generating 3D neuron segmentation and reconstruction results.

3. Results

In this section, we first present the experimental setting and evaluation methods. Next, we evaluate the proposed method and compare it with other state-of-the-art methods on two public data sets of EM structure segmentation tasks, namely, the ISBI EM segmentation challenge [8] and the mouse piriform cortex data set [32]. The results on both data sets show the superiority and robustness of our method. In addition, we apply the proposed method to the Mouse ATUM-SEM data set to demonstrate the effectiveness of RC-SPCNet in handling large-scale data.

3.1. Implementation details

The proposed network is implemented on a NVIDIA Quadro RTX 6000 GPU machine with Keras deep learning framework [33] and TensorFlow backend. In the training process, the training set is further divided into training data and validation data, and the ratio of training data to validation data is set to 7:3. The network is optimized using the Adam optimizer [34] with initial learning rate of 0.0001, the total epoch and the batch size are set to 50 and 2, respectively; then, we reduced the learning rate when the validation loss stopped decreasing, the reduction factor and patience are set to 0.7 and 3, respectively. To improve the efficiency of training, an early stopping strategy is adopted to stop training when the monitored quantity stopped improving, and the patience is set to 5 epochs. We chose the pixel-wise mean squared error as the loss function of the network. The codes are available online[2].

In the lifted multicut post-processing, we used the vigra image processing library [35] for image manipulation, the nifty (https://github.com/DerThorsten/nifty) and elf library (https://github.com/constantinpape/elf) for the graph-based segmentation and solving discrete optimization problems, the scikit-learn library [36] for random forest, and napari [37] for visualization. In the experiment of ISBI and Mouse Cortex datasets, we used all the features in Ref. [16] to obtain promising results. In the experiment of Mouse ATUM-SEM dataset, we only extracted the features of regular edges and the connectedness feature and the region feature of lifted edges to improve the speed of the algorithm.

3.2. Evaluation methods

In the evaluation stage, the probability map of the neuronal boundaries is first transformed into a binary image via different post-processing methods; then, the binary result is evaluated based on the topological metrics in the ISBI EM Segmentation Challenge [9]. The metrics include maximal foreground-restricted Rand score after thinning ($V^{Rand}_{thinned}$) and maximal foreground-restricted information theoretic score after thinning ($V^{info}_{thinned}$).

The $V^{Rand}_{thinned}$ index is a special normalized version of the Rand F-score [38] that is mainly used to evaluate the accuracy of structural correlation between a prediction and the ground truth. The higher the index is, the more similar the prediction and ground truth are. $V^{info}_{thinned}$ is a variation of the information metric used to calculate the similarity between a prediction and the ground truth [39]. According to Ref[9]., $V^{Rand}_{thinned}$ is more robust than $V^{info}_{thinned}$; thus, the ranking is sorted with respect to the former metric.

[2] https://github.com/xiaochiHNU/RC_SPCNet

For the ISBI EM segmentation challenge, we submitted predictions to the challenge website and obtained the evaluation results for two metrics. For the mouse piriform cortex segmentation, we used a script in ImageJ [40] to obtain the evaluation results, which calculated the best $V^{Rand}_{thinned}$ score over a set of threshold values of the prediction results.

3.3. ISBI EM Segmentation challenge

The public anisotropic data set was collected from the ventral nerve line of Drosophila melanogaster and imaged by ssTEM (the voxel resolution is 4 nm $\times$ 4 nm $\times$ 50 nm) and is divided into a training set and test set. The training set consists of 30 original images with a size of 512 $\times$ 512 and the corresponding ground truth, and the test set contains another batch of 512 $\times$ 512 $\times$ 30 images with the ground truth that is used by the organizers for evaluation. The challenge has attracted many well-known international universities to participate (including Heidelberg University, Princeton University, ETH Zõrich, and Allen Institute). Although the challenge finished in 2012, the submission system has remained open, and the leader board is constantly updated. As of April 2021, 243 groups have participated in the challenge, and there have been over 8000 submissions.

Since the number of EM images in the training set is small, a data augmentation strategy is adopted to address the problem of small training data sets in deep networks. In this process, we first applied image elastic distortion [41] to the entire training stack by warping all image pixels (both raw data and the corresponding labels). The parameters of the elastic distortion (scaling factor α and standard variance of Gaussian function σ) are set to 2500 and 50, respectively. Next, rotation and flipping are applied to the enriched data set. There are 8 transformation variants, which consisted of rotations of 0°, 90°, 180°, and 270° and the horizontal flipping over the xy-plane. The training set contains 9000 samples (size of 512 $\times$ 512) after the data augmentation, which is conducive to train the proposed network. In the inference process, we used the TTA strategy to further improve the performance. Additionally, we utilized cascaded random forest, ResNet50 and ResNet38 as the baselines and discuss several possible variations for the proposed network and post-processing strategies.

Figs. 5 and 6 present qualitative comparisons of the baselines and different variations of our method. In Fig. 5, the residual contextual network (RCNet) is the variation of the proposed network with classic deconvolutional layers. The segmentation results of RCNet are more accurate than those of ResNet50 and ResNet38 (see the red arrows), which is mainly attributed to the utilization of multilevel contextual cues. Since different regions may depend on different receptive field sizes, a network with multilevel contextual information can effectively recognize neuronal boundaries and other intracellular structures. Compared with the above methods, RC-SPCNet (k=2) provides fewer false positives and more accurate segmentation results since SPC can restore the details of the image that might be missed by deconvolution. In Fig. 6, the proposed network with TTA yields the best results. Although the element-wise maximum strategy increases the noise and makes the predicted boundaries thicker (see the blue arrows), it also ensures continuity of the boundaries and improves the accuracy of the neuronal structure segmentation. Finally, the lifted multicut approach refines the discontinuous boundaries of the probability maps (see the green arrows). The quantitative results of the variations of our network are summarized in Table 1. First, RCNet performed better than cascaded random forest, ResNet50 and ResNet38. Since the proposed networks include contextual skip connection, it is essential to determine whether the segmentation accuracy could be improved by adding further skip connections. In our work, we compared the performance between networks with zero, one and two skip con-

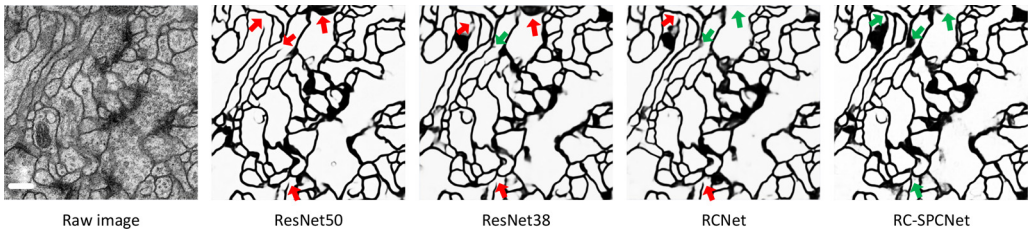

Fig. 5. Qualitative comparison between different segmentation methods. Left to right: Raw image of slice 11/30 in the ISBI EM segmentation challenge test set; Resnet50 result; Resnet38 result; RCNet result with skip connections ($k = 2$); RC-SPCNet result with skip connections ($k = 2$). Red and green arrows denote poor and good results, respectively. Scale bar: 200 nm.

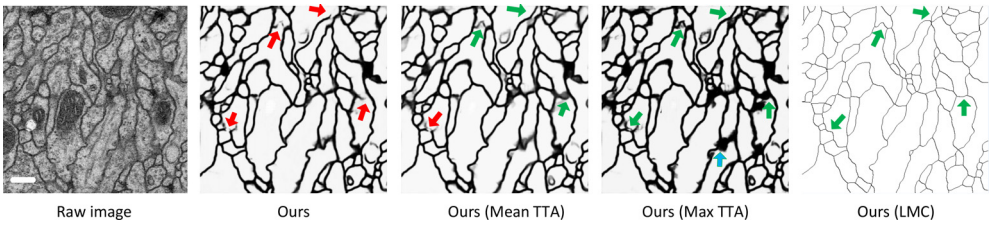

Fig. 6. Qualitative comparison using different post-processing methods. "Ours" denotes the RC-SPCNet results. Left to right: raw images; our results; our results with the mean value of TTA; our results with the maximum value of TTA; our results with the lifted multicut post-processing. Scale bar: 200 nm.

Table 1
Quantitative performance on the ISBI EM segmentation challenge. THR denotes thresholding, k denotes the number of skip connections, Mean and Max TTA represent the mean and maximum values of the test-time augmentation with 8 variants, respectively. RF represents random forest. The overall best scores are in bold.

Methods	post-processing	$V^{Rand}_{thinned}$	$V^{info}_{thinned}$
Cascaded RF	THR	0.893902298	0.954403639
Resnet50	THR	0.971685325	0.982677218
Resnet38	THR	0.973820341	0.987658760
RCNet (k=0)	THR	0.974200530	0.987571105
RCNet (k=1)	THR	0.975506699	0.987210826
RCNet (k=2)	THR	0.977312410	0.987823374
RCNet (k=2)	Mean TTA	0.977713645	0.988540642
RCNet (k=2)	Max TTA	0.983867935	0.990189890
RCNet (k=2)	Multicut	0.983563573	0.990630782
RC-SPCNet (k=2)	Mean TTA	0.981484972	0.988040074
RC-SPCNet (k=2)	Max TTA	0.984264734	0.989370141
RC-SPCNet (k=2)	Lifted multicut	**0.987877739**	**0.990920188**

Table 2
The top 10 groups of the ISBI EM segmentation challenge on neuronal structures. For more details about the leader board, please refer to the website (http://brainiac2.mit.edu/isbi_challenge/leaders-board-new).

Group name	$V^{Rand}_{thinned}$	$V^{info}_{thinned}$	Rank
* human values **	0.997847778	0.998997659	-
MDC Berlin PatchPerPix	0.988290649	0.991641507	1
IAL MutexWS	0.987922250	0.991833594	2
CASIA-MIRA (Ours)	**0.987877739**	**0.990920188**	**3**
IAL - Steerable Filter CNN	0.986800916	0.991438892	4
Lances	0.986106874	0.988456542	5
ACE-Net [42]	0.985032746	0.989490497	6
M2FCN-MFA [15]	0.983834543	0.989805687	7
HO-CRF	0.983654610	0.989568936	8
HVCL@UNIST [13]	0.983651122	0.991303595	9
Afshin Khadangi	0.983599730	0.987164611	10
RC-SPCNet-TTA (Ours)	0.984264734	0.989370141	

A total of 243 teams participated in the ISBI EM Segmentation challenge as of April 6th, 2021.

nections, as shown in Table 1. The $V^{Rand}_{thinned}$ index of RCNet (k=1) and RCNet (k=2) increased 0.13% and 0.18% compared with that of RCNet (k=0), respectively, which demonstrates that our multi-context model is beneficial for neuronal structure segmentation. For the post-processing and TTA schemes, the improvement in the mean TTA is constrained. However, the maximum TTA and multicut achieved 0.66% and 0.63% improvements in V_{rand} compared with that of the threshold method. In addition, compared to the deconvolution strategy, RC-SPCNet led to a considerable performance improvement.

Table 2 presents the top 10 teams of the ISBI EM segmentation challenge on neuronal structures. Our method is among the top of the leader board and yields 0.98788 V_{rand} and 0.99092 V_{info}, which are close to the accuracy of human experts and surpass the performance of most state-of-the-art methods. Note that some promising methods apply 3D networks or complicated post-processing to enhance performance, such as IAL MutexWS and DerThorsten [16]. Our network with simple post-processing (RC-SPCNet-TTA) also achieved a V_{rand} score of 0.9842, which is comparable with that of these promising methods.

3.4. Mouse piriform cortex data set

The public anisotropic ssTEM data set[3] from the piriform cortex of mouse was acquired by the Seung laboratory and annotated by Kisuk Lee and Ashwin Vishwanathan [32]. The data set consists of 4 anisotropic EM stacks with voxel resolution of nearly 7 nm × 7 nm × 40 nm. Following [43], we adopted stack1 and stack2 as the training data and stack3 and stack4 for testing. The data augmentations of the experiment include image elastic distortion, cropping, flipping, and rotating by 90°, 180°, and 270°. After the data augmentation, the training set includes nearly 7000 samples (size of 256 × 256).

Fig. 7 presents the qualitative results from the ground truth and variations of the RC-SPCNet models, where the number of skip connections is set up to 2. The false segmentation of ultrastructures, such as mitochondria and vesicles, can be reduced by using multilevel contextual information. SPC and lifted multicut are beneficial for refining the discontinuous boundaries and removing false neuronal boundaries (indicated by green arrows).

Table 3 shows a qualitative comparison between RC-SPCNet and other promising methods. As mentioned in [43], SegNet [44], PSP-

[3] http://seunglab.org/data/

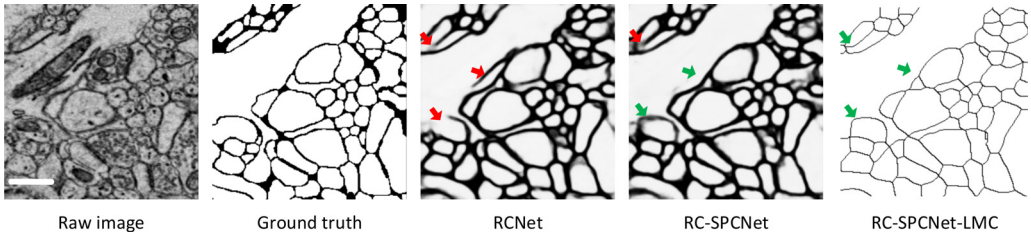

Fig. 7. Qualitative comparison on the mouse piriform cortex data set. Scale bar: 500 nm.

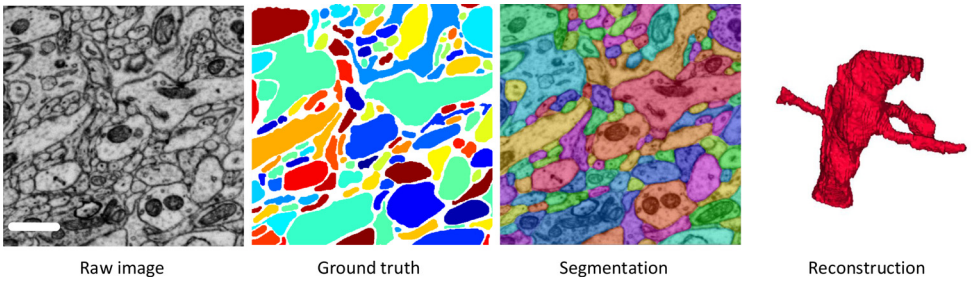

Fig. 8. Our segmentation and reconstruction results in the mouse piriform cortex testing data set. In the ground truth and segmentation results, different colors indicate different labels. Scale bar: 800 nm.

Table 3
Quantitative comparison between different segmentation networks on the mouse piriform cortex data set based on $V^{Rand}_{thinned}$. The overall best scores are in bold.

Methods	Stack3	Stack4
SegNet [44]	0.8164	0.6399
PSPNet [45]	0.7367	0.8404
LinkNet [46]	0.8201	0.8100
GCN [47]	0.8286	0.8583
LKRFNet [43]	0.8534	0.8685
RCNet	0.9239	0.9232
RC-SPCNet	0.9440	0.9245
RC-SPCNet (Focal-loss)	0.9556	**0.9323**
RC-SPCNet-LMC	**0.9562**	0.9318

Net [45], LinkNet [46] and GCN [47] were proposed to solve natural image segmentation problems and might ignore the characteristics of EM images and achieve unsatisfactory performance in neuronal structure segmentation. In Ref[43]., the large kernel refine fusion network was proposed to segment the neuron effectively, and post-processing of the refined boundary was used to obtain better results. However, the shape of neuronal boundaries is always narrow and long, and large kernel convolution might cause a loss of detailed information. As shown in Table 3, our method, RC-SPCNet, achieves $V^{Rand}_{thinned}$ scores of 0.9440 and 0.9245 in different testing stacks, which far surpass those of other methods, and the $V^{Rand}_{thinned}$ scores of the best proposed model with the use of lifted multicut are 0.9562 and 0.9318. Additionally, considering that the ratio of the neuronal boundary to non-neuronal boundary is nearly 1:4, we adopted focal loss [48] to address this class-imbalance problem. The results show that RC-SPCNet with focal loss yields $V^{Rand}_{thinned}$ scores of 0.9556 and 0.9323, which are close to the results of RC-SPCNet with the lifted multicut post-processing.

Next, we present a comparison between our segmentation and the corresponding ground truth in Fig. 8. Furthermore, the segmentation results are imported into ImageJ to show the 3D reconstruction of a neurite. As illustrated in Fig. 8, the segmentation and the corresponding ground truth are very similar, and the reconstructions of dendrites and dendritic spines are intact and authentic, which confirms the validity of our method.

3.5. Mouse ATUM-SEM data set

Sections 3.3 and 3.4 fully illustrate the effectiveness of the proposed method. Subsequently, we applied the proposed RC-SPCNet to the larger-scale Mouse ATUM-SEM data set and obtained the segmentation and reconstruction of neurons. The whole Mouse ATUM-SEM data set (volume of 15.2 × 17.2 × 8.9 um^3) from the mouse cortex was acquired by the Institute of Neuroscience and Institute of Automation and consists of 178 sections with thicknesses of approximately 50 nm and 2 nm pixel size in the xy plane. We adopted 30 stack images (size of 2048 × 2,048) as the training data set, and the corresponding ground truth masks are annotated by several experts using ImageJ with the TrakEM plug-in. The data augmentation and training strategy is similar to that in the above Section 3.4. After the data augmentation, the training set consists of nearly 8000 samples (size of 512 × 512). Considering that the volume of the mouse ATUM-SEM data set is relatively large, we utilized the hierarchical solver [31] to improve the running speed of the lifted multicut. The segmentation of neuron boundaries and structure are shown in Fig. 9. The neuronal boundaries are recognized precisely, and most of the segmentation results of the cell body, dendrites and axons are correct.

Next, we imported the segmentation results into ImageJ and finished the validation of most of the neuronal cell body, dendrites and axons in a few hours. Some 3D reconstruction results are presented in Fig. 9: the dendrites and axons are complete and continuous, which demonstrates the feasibility of our method.

3.6. Computational costs

In this subsection, we assess the time consumed by the complete manual annotation and the proposed method to demonstrate the superiority of our automated method.

The comparisons of computational times between manual annotation and the proposed method are shown in Table 4. On average, it takes nearly 320 min to manual annotate all the neuron in a 512 × 512 × 30 image stack using ImageJ with TrakEM plug-in. The computational costs of the proposed method consist of training and inference, the part of inference can be further divided into the network prediction and the post-processing, and the running time of the lifted multicut post-processing is far more than that of

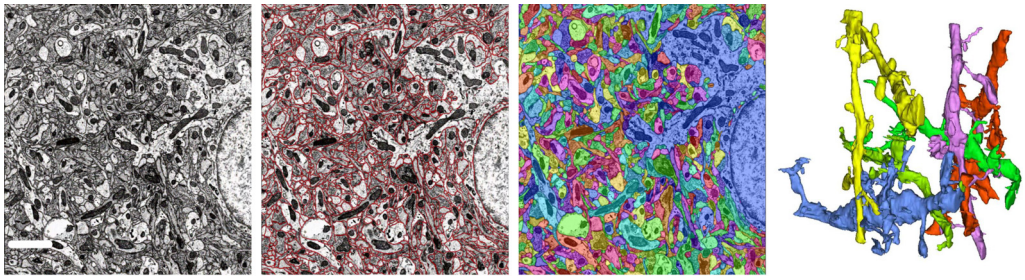

Fig. 9. Mouse ATUM-SEM data set segmentation results. Scale bar: 4 um.

Table 4
Computational costs comparison between manual annotation and the proposed method.

Data set	Volume size	Manual annotation	Our proposed method		
			Training	Inference	Total
ISBI	512 × 512 × 30	320 mins	176 mins	60 min	236 mins
Mouse Cortex	512 × 512 × 169 256 × 256 × 121	2,135 mins	153 mins	210 mins	363 mins
Mouse ATUM-SEM	7,600 × 8600 × 178	473,400 mins	182 mins	630 mins	812 mins

the network prediction. For the ISBI data set, the total computational time of our method is 236 min, which is close to the time consumed of manual annotation. For the larger data sets (mouse cortex and mouse ATUM-SEM), the proposed method only takes 363 min and 812 min, respectively. However, the estimated manual annotation time is 2135 min and 473,400 min.

All the costs are measured on the workstation with Quadro RTX 6000 GPU and Intel Xeon Gold 6234 CPU. From the comparison, it is important to note that the proposed method reduces the segmentation costs by orders of magnitude in large-scale data. Although the accuracy of manual annotation by experts is still better than that of automatic reconstruction algorithm, the experimental results illustrate that our method is capable of reducing the manual workload and assisting neuroanatomists in neuronal structure segmentation.

4. Discussion

Currently, the mainstream research methods for microconnectomics research reconstruct the connectome using electron microscopy. Neuronal structure segmentation in EM images is one of the greatest challenges due to the various shapes and sizes of neurons, image deformation and ambiguous intracellular structures. In this article, we proposed an effective and automatic neuronal structure segmentation method for anisotropic EM data and reduce the burden of manual neurite validation.

As mentioned above, our network outperforms most promising approaches on two public data sets and yields state-of-the-art performance. For the ISBI EM segmentation challenge, we submitted the predicted results to the challenge website and compared our results with those of 242 other teams: our best result now ranks in the top 3 and achieves a Rand score of 0.98788, which is close to the accuracy of human experts. For the mouse piriform cortex data set, we compared our results with baselines of [43–47]. As shown in Table 3, our best model yields Rand scores of 0.9562 and 0.9318, respectively, which are much better than those of the baseline networks, thereby demonstrating the superiority of our method in terms of neuronal structure segmentation. Additionally, we utilized the proposed method to obtain the neuronal structure segmentation and reconstruction results in large-scale EM images.

The promising performance of our network can be attributed to a variety of factors. First, the proposed deep and spatially efficient residual network is beneficial to extracting the features of neuronal boundaries and avoiding the vanishing gradient problem. Additionally, the incorporation of multiscale contextual cues makes it possible to increase the contextual information and distinguish the neuronal boundaries and other ultrastructural objects effectively. More importantly, the subpixel convolutional layer is used to upsample feature maps, which helps recover information about the neuronal boundary and further improve the accuracy of segmentation.

Recently, methods based on 3D networks and their variants can effectively utilize spatial information to produce more accurate results in 3D neuron segmentation. On this basis, Lee et al. proposed a variant of 3D U-Net to predict affinities and long-range affinities between nearest neighbor voxels, which surpassed the human accuracy in SNEMI 3D challenge [49]. Januszewski et al. proposed flood-filled network for automated neuronal structure reconstruction, the framework contains a 3D recurrent neural network that iteratively optimizes the neuron segmentation results [50]. Lee et al. adopted a 3D deep metric learning network to learn dense voxel embeddings for highly accurate 3D neuron segmentation and reconstruction [51]. In this work, we proposed RC-SPCNet to segment the neuronal boundaries and then utilized lifted multicut to obtain the 3D connection of neuron structures. The experimental results on the ISBI EM segmentation challenge and mouse piriform cortex data set proved that our network achieved promising results. For the evaluation of the 3D reconstructions, we submitted our predictions to the SNEMI 3D challenge, the Rand error of the proposed method is 0.0801, though not as good as the results in Ref. [49–51], but is not far away from the human values of 0.05998.

Considering that we only chose a 2D network to obtain the neuronal boundaries, the proposed network could not make full use of the 3D information in EM stacks. Additionally, in the post-processing of lifted multicut, the weights of the lifted edges are calculated by random forest classifier after a long time training, which increase the time consumption and complexity of the algorithm. The future works are as follows: 1. Concentrate on modifying the RC-SPCNet to 3D neural network and implementing it on 3D neuron structures segmentation in both anisotropic and isotropic data sets. 2. Integrate the calculation of the lifted edge into the neural network to reduce the time consumption and further improve the performance of the post-processing.

In conclusion, this article proposes a neural structure segmentation and reconstruction pipeline based on a deep residual contex-

tual and subpixel convolution network that incorporates multilevel contextual information and subpixel convolution to obtain neuronal boundaries. Furthermore, lifted multicut is used as a post-processing step to refine the segmentation results and obtain the 3D reconstruction of neurons. Our method currently ranks among the top performers in the ISBI EM segmentation challenge and yields the best performance results in the mouse piriform cortex data set. Based on this method, we reconstructed the mouse ATUM-SEM data set ($15.2 \times 17.2 \times 8.9$ um^3) and obtained most of the neuronal structures. The experimental results confirmed that our automated method achieved state-of-the-art performance and provides neuroanatomists with an effective approach to obtain the neuronal structure segmentation and reduce their burden with respect to validation.

Funding statement

This work is supported by the Science and Technology Innovation 2030 Major Projects of China (Grant No. 2021ZD0204503), the National Natural Science Foundation of China (Grant No. 61673381 and 32171461), the Education Department of Hainan Province (Grant No. Hnky2021-13), the Hainan Natural Science Foundation (Grant No. 621QN216), the International Partnership Program of Chinese Academy of Science (Grant No. 153D31KYSB20170059), the Strategic Priority Research Program of Chinese Academy of Science (Grant No. XDA16021104 and XDB32030208), and the Program of Beijing Municipal Science and Technology Commission (Grant No. Z201100008420004).

Declaration of Competing Interest

The authors declare that they have no competing interests.

The authors declared that they have no conflict of interest and the paper presents their own work which does not been infringe any third-party rights, especially authorship of any part of the article is an original contribution, not published before and not being under consideration for publication elsewhere. The order of authors listed in the manuscript has been approved by both authors.

References

[1] J.W. Lichtman, W. Denk, The big and the small: challenges of imaging the brain circuits, Science 334 (6056) (2011) 618–623.
[2] M. Helmstaedter, P.P. Mitra, Computational methods and challenges for large-scale circuit mapping, Curr. Opin. Neurobiol. 22 (1) (2012) 162–169.
[3] M. Helmstaedter, Cellular-resolution connectomics: challenges of dense neural circuit reconstruction, Nat. Methods 10 (6) (2013) 501.
[4] Q. Xie, X. Chen, L. Shen, G. Li, H. Ma, H. Han, Micro reconstruction system for brain, Syst. Eng.-Theory Pract. 37 (11) (2017) 3006–3017.
[5] S.-y. Takemura, A. Bharioke, Z. Lu, A. Nern, S. Vitaladevuni, P.K. Rivlin, W.T. Katz, D.J. Olbris, S.M. Plaza, P. Winston, et al., A visual motion detection circuit suggested by drosophila connectomics, Nature 500 (7461) (2013) 175.
[6] K.M. Harris, E. Perry, J. Bourne, M. Feinberg, L. Ostroff, J. Hurlburt, Uniform serial sectioning for transmission electron microscopy, J. Neurosci. 26 (47) (2006) 12101–12103.
[7] J. Cepelewicz, The US government launches a $100-million apollo project of the brain, Sci. Am. (2016).
[8] ISBI, Segmentation of neuronal structures in em stacks challenge, 2012.
[9] I. Arganda-Carreras, S.C. Turaga, D.R. Berger, D. Cireşan, A. Giusti, L.M. Gambardella, J. Schmidhuber, D. Laptev, S. Dwivedi, J.M. Buhmann, et al., Crowdsourcing the creation of image segmentation algorithms for connectomics, Front. Neuroanat. 9 (2015) 142.
[10] 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.
[11] M.F. Stollenga, W. Byeon, M. Liwicki, J. Schmidhuber, Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation, in: Advances in Neural Information Processing Systems, 2015, pp. 2998–3006.
[12] 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.
[13] T.M. Quan, D.G. Hilderbrand, W.-K. Jeong, FusionNet: a deep fully residual convolutional neural network for image segmentation in connectomics, arXiv preprint arXiv:1612.05360 (2016).
[14] M. Drozdzal, G. Chartrand, E. Vorontsov, M. Shakeri, L. Di Jorio, A. Tang, A. Romero, Y. Bengio, C. Pal, S. Kadoury, Learning normalized inputs for iterative estimation in medical image segmentation, Med. Image Anal. 44 (2018) 1–13.
[15] W. Shen, B. Wang, Y. Jiang, Y. Wang, A. Yuille, Multi-stage multi-recursive-input fully convolutional networks for neuronal boundary detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2391–2400.
[16] T. Beier, C. Pape, N. Rahaman, T. Prange, S. Berg, D.D. Bock, A. Cardona, G.W. Knott, S.M. Plaza, L.K. Scheffer, et al., Multicut brings automated neurite segmentation closer to human performance, Nat. Methods 14 (2) (2017) 101–102.
[17] M. Vidotto, E.D. Momi, M. Gazzara, L.S. Mattos, G. Ferrigno, S. Moccia, FC-NN-based axon segmentation for convection-enhanced delivery optimization, Int. J. Comput. Assist. Radiol. Surg. 14 (3) (2019) 493–499.
[18] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: International Conference on Neural Information Processing Systems, 2012.
[19] X. Hao, Y. Du, H. Yu, Y. Chang, Z. Xu, Y. Tang, Open set face recognition with deep transfer learning and extreme value statistics, Int. J. Wavelets Multiresolution Inf. Process. (7) (2018). S0219691318500340
[20] C. Zhu, X.C. Yin, Effective human detection via multi-model classification and adaptive late fusion, Int. J. Wavelets Multiresolution Inf. Process. 16 (2) (2018).
[21] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[22] A. Veit, M.J. Wilber, S. Belongie, Residual networks behave like ensembles of relatively shallow networks, in: Advances in Neural Information Processing Systems, 2016, pp. 550–558.
[23] Z. Wu, C. Shen, A.v. d. Hengel, Wider or deeper: revisiting the resnet model for visual recognition, arXiv preprint arXiv:1611.10080 (2016).
[24] W. Shi, J. Caballero, F. Huszár, J. Totz, A.P. Aitken, R. Bishop, D. Rueckert, Z. Wang, Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1874–1883.
[25] P. Wang, P. Chen, Y. Yuan, D. Liu, Z. Huang, X. Hou, G. Cottrell, Understanding convolution for semantic segmentation, in: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2018, pp. 1451–1460.
[26] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, arXiv preprint arXiv:1502.03167 (2015).
[27] D.-A. Clevert, T. Unterthiner, S. Hochreiter, Fast and accurate deep network learning by exponential linear units (ELUs), arXiv preprint arXiv:1511.07289 (2015).
[28] L.N. Smith, N. Topin, Deep convolutional neural network design patterns, arXiv preprint arXiv:1611.00847 (2016).
[29] H. Noh, S. Hong, B. Han, Learning deconvolution network for semantic segmentation, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1520–1528.
[30] T. Beier, B. Andres, U. Köthe, F.A. Hamprecht, An efficient fusion move algorithm for the minimum cost lifted multicut problem, in: European Conference on Computer Vision, Springer, 2016, pp. 715–730.
[31] C. Pape, A. Matskevych, A. Wolny, J. Hennies, G. Mizzon, M. Louveaux, J. Musser, A. Maizel, D. Arendt, A. Kreshuk, Leveraging domain knowledge to improve microscopy image segmentation with lifted multicuts, Front. Comput. Sci. 1 (2019).
[32] K. Lee, A. Zlateski, V. Ashwin, H.S. Seung, Recursive training of 2D-3D convolutional networks for neuronal boundary prediction, in: Advances in Neural Information Processing Systems, 2015, pp. 3573–3581.
[33] F. Chollet, et al., Keras. https://keras.io.
[34] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).
[35] U. Köthe, The VIGRA image analysis library, University of Heidelberg, Heidelberg, Germany, 2013.
[36] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[37] N. Contributors, napari: A Multi-Dimensional Image Viewer for Python, Zenodo, 2019, doi:10.5281/zenodo. 3555620
[38] R. Unnikrishnan, C. Pantofaru, M. Hebert, Toward objective evaluation of image segmentation algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 29 (6) (2007) 929–944.
[39] M. Meilă, Comparing clusterings: an axiomatic view, in: Proceedings of the 22nd International Conference on Machine Learning, ACM, 2005, pp. 577–584.
[40] B. Schmid, J. Schindelin, A. Cardona, M. Longair, M. Heisenberg, A high-level 3D visualization API for java and imagej, BMC Bioinformatics 11 (1) (2010) 274.
[41] P.Y. Simard, D. Steinkraus, J.C. Platt, Best practices for convolutional neural networks applied to visual document analysis, in: null, IEEE, 2003, p. 958.
[42] Y. Zhu, Z. Chen, S. Zhao, H. Xie, W. Guo, Y. Zhang, ACE-Net: biomedical image segmentation with augmented contracting and expansive paths, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 712–720.
[43] D. Liu, D. Zhang, Y. Song, C. Zhang, H. Huang, M. Chen, W. Cai, Large kernel refine fusion net for neuron membrane segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 2212–2220.

[44] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: a deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.
[45] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid scene parsing network, in: IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2881–2890.
[46] A. Chaurasia, E. Culurciello, LinkNet: exploiting encoder representations for efficient semantic segmentation, in: Visual Communications and Image Processing (VCIP), 2017 IEEE, IEEE, 2017, pp. 1–4.
[47] C. Peng, X. Zhang, G. Yu, G. Luo, J. Sun, Large kernel matters – improve semantic segmentation by global convolutional network, in: Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on, IEEE, 2017, pp. 1743–1751.
[48] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2980–2988.
[49] K. Lee, J. Zung, P. Li, V. Jain, H.S. Seung, Superhuman accuracy on the SNEMI3D connectomics challenge, arXiv preprint arXiv:1706.00120 (2017).
[50] M. Januszewski, J. Kornfeld, P.H. Li, A. Pope, T. Blakely, L. Lindsey, J. Maitin-Shepard, M. Tyka, W. Denk, V. Jain, High-precision automated reconstruction of neurons with flood-filling networks, Nat. Methods 15 (8) (2018) 605–610.
[51] K. Lee, R. Lu, K. Luther, H.S. Seung, Learning and segmenting dense voxel embeddings for 3D neuron reconstruction, IEEE Trans. Med. Imaging 40 (12) (2021) 3801–3811.