


METHODOLOGY ARTICLE

Open Access



Effective automated pipeline for 3D reconstruction of synapses based on deep learning

Chi Xiao^{1,2}, Weifu Li^{1,3}, Hao Deng⁴, Xi Chen¹, Yang Yang^{5,6}, Qiwei Xie^{1,7*}  and Hua Han^{1,2,6*}

Abstract

Background: The locations and shapes of synapses are important in reconstructing connectomes and analyzing synaptic plasticity. However, current synapse detection and segmentation methods are still not adequate for accurately acquiring the synaptic connectivity, and they cannot effectively alleviate the burden of synapse validation.

Results: We propose a fully automated method that relies on deep learning to realize the 3D reconstruction of synapses in electron microscopy (EM) images. The proposed method consists of three main parts: (1) training and employing the faster region convolutional neural networks (R-CNN) algorithm to detect synapses, (2) using the z-continuity of synapses to reduce false positives, and (3) combining the Dijkstra algorithm with the GrabCut algorithm to obtain the segmentation of synaptic clefts. Experimental results were validated by manual tracking, and the effectiveness of our proposed method was demonstrated. The experimental results in anisotropic and isotropic EM volumes demonstrate the effectiveness of our algorithm, and the average precision of our detection (92.8% in anisotropy, 93.5% in isotropy) and segmentation (88.6% in anisotropy, 93.0% in isotropy) suggests that our method achieves state-of-the-art results.

Conclusions: Our fully automated approach contributes to the development of neuroscience, providing neurologists with a rapid approach for obtaining rich synaptic statistics.

Keywords: Electron microscope, Synapse detection, Deep learning, Synapse segmentation, 3D Reconstruction of synapses

Background

A synapse is a structure that permits a neuron (or nerve cell) to pass an electrical or chemical signal to another neuron, and it has an important responsibility in the neural system. If we consider the brain network to be a map of connections, then neurons and synapses can be considered as the dots and lines, respectively, and it can be hypothesized that the synapse is one of the key factors for researching connectomes [1–3]. In addition, synaptic plasticity is associated with learning and memory. Sensory experience, motor learning and aging are found to induce

alterations in presynaptic axon boutons and postsynaptic dendritic spines [4–6]. Consequently, understanding the mechanism of synaptic plasticity will be conducive to the prevention and treatment of brain diseases. To study the correlation between synaptic growth and plasticity and to reconstruct neuronal connections, it is necessary to obtain the number, location and structure of synapses in neurons.

According to the classification of synaptic nerve impulses, there are two types of synapses: chemical synapses and electrical synapses. In this study, we focus on the chemical synapse, which consists of presynaptic (axonal) membrane, postsynaptic (dendritic) membrane and a 30–60 nm synaptic cleft. Because of its limited resolution, optical microscopy cannot provide sufficient resolution to reveal these fine structures. Fortunately, it is now possible to more closely examine the synapse structure

*Correspondence: qiwei.xie@ia.ac.cn; hua.han@ia.ac.cn

¹Institute of Automation, Chinese Academy of Sciences, 95 Zhongguancun East Road, 100190 Beijing, China

⁷Data Mining Lab, Beijing University of Technology, 100 Ping Le Yuan, 100124 Beijing, China

Full list of author information is available at the end of the article



due to the rapid development of electron microscopy (EM). In particular, focused ion beam scanning electron microscopy (FIB-SEM) [7] can provide nearly 5 nm imaging resolution, which is conducive to obtaining the very fine details of ultrastructural objects; however, this technique is either limited to a small section size ($0.1\text{ mm} \times 0.1\text{ mm}$) or provides blurred imaging.

By contrast, automated tape-collecting ultramicrotome scanning electron microscopy (ATUM-SEM) [8] offers anisotropic voxels with a lower imaging resolution in the z direction ($2\text{ nm} \times 2\text{ nm} \times 50\text{ nm}$), but it is capable of working with large-area sections ($2.5\text{ mm} \times 6\text{ mm}$). Moreover, ATUM-SEM does not damage any sections; thus, the preserved sections can be imaged and analyzed many times. Considering volume and resolution, this paper employs ATUM-SEM and FIB-SEM image stacks to verify the validity and feasibility of our algorithms.

Note that EM images with higher resolution will inevitably produce more data in the same volume; thus, synapse validation requires a vast amount of laborious and repetitive manual work. Consequently, an automated synapse reconstruction pipeline is essential for analyzing large volumes of brain tissue [9]. Prior works on synapse detection and segmentation investigated a range of approaches. Mishchenko et al. [10] developed a synaptic cleft recognition algorithm to detect postsynaptic densities in serial block-face scanning electron microscopy (SBEM) [11] image stacks. However, this method was effective for synapse detection only if the prior neuron segmentation was satisfactory. Navlakha et al. [12] presented an original experimental technique for selectively staining synapses, and then they utilized a semi-supervised method to train classifiers such as support vector machine (SVM), AdaBoost and random forest to identify synapses. Similarly, Jagadeesh et al. [13] presented a new method for synapse detection and localization. This method first characterized synaptic junctions as ribbons, vesicles and clefts, and then it utilized maximally stable extremal region (MSER) to design a detector to locate synapses. However, all these works [10, 12, 13] ignored the contextual information of synapses.

For the above reasons, Kreshuk et al. [14] presented a contextual approach for automated synapse detection and segmentation in FIB-SEM image stacks. This approach adopted 35 appearance features, such as magnitude of Gaussian gradient, Laplacian of Gaussian, Hessian matrix and structure tensor, and then it employed a random forest classifier to produce synapse probability maps. Nevertheless, this approach neglected the asymmetric information produced by the presynaptic and postsynaptic regions, which led to some inaccurate results. Becker et al. [15] utilized contextual information and different Gaussian kernel functions to calculate synaptic characteristics, and then they employed these features to train an

AdaBoost classifier to obtain synaptic clefts in FIB-SEM image stacks. Similarly, Kreshuk et al. [16] proposed an automated approach for synapse segmentation in serial section transmission electron microscopy (ssTEM) [17] image stacks. The main idea was to classify synapses from 3D features and then segment synapses by using the Ising model and object-level features classifier. Ref. [16] did not require prior segmentation and achieved a good error rate. Sun et al. [18] focused on synapse reconstruction in anisotropic image stacks, which were acquired through ATUM-SEM; detected synapses with cascade AdaBoost; and then utilized continuity to delete false positives. Subsequently, the variational region growing [19] was adopted to segment synaptic clefts. However, the detection accuracies of Ref. [16] and Ref. [18] were not satisfactory, and the segmentation results lacked smoothness.

Deep neural networks (DNNs) have recently been widely applied in solving medical imaging detection and segmentation problems [20–23] due to their extraordinary performance. Thus, the application of DNNs to synapse detection in EM data holds great promise. Roncal et al. [24] proposed a deep learning classifier (VESICLE-CNN) to segment synapses directly from EM data without any prior knowledge of the synapse. Staffler et al. [25] presented SynEM, which focused on classifying borders between neuronal processes as synaptic or non-synaptic and relied on prior neuron segmentation. Dorkenwald et al. [26] developed the SyConn framework, which used deep learning networks and random forest classifiers to obtain the connectivity of synapses.

In this paper, we introduce a fully automated method for realizing the 3D dense reconstruction of synapses in FIB-SEM and ATUM-SEM images by combining a series of effective detection and segmentation methods. The image datasets are depicted in Fig. 1. To avoid false distinctions between a synaptic cleft and membrane, we utilize contextual information to consider the presynaptic membrane, synaptic cleft and postsynaptic membrane as a whole, and then we adopt a deep learning detector [27] to obtain the accurate localization of synapses. Subsequently, a screening method with z -continuity is proposed to improve the detection precision. To precisely segment synapses, the Dijkstra algorithm [28] is employed to obtain the optimal path of the synaptic cleft, and the GrabCut algorithm [29] is applied for further segmentation. Finally, we utilize ImageJ [30] to visualize the 3D structure of synaptic clefts, and we compare our results with other promising results obtained by Refs. [15, 18, 19, 23]. By using deep learning, z -continuity and GrabCut, our approach performs significantly better than these methods.

Method

The proposed automated synapse reconstruction method for EM serial sections of biological tissues can be divided

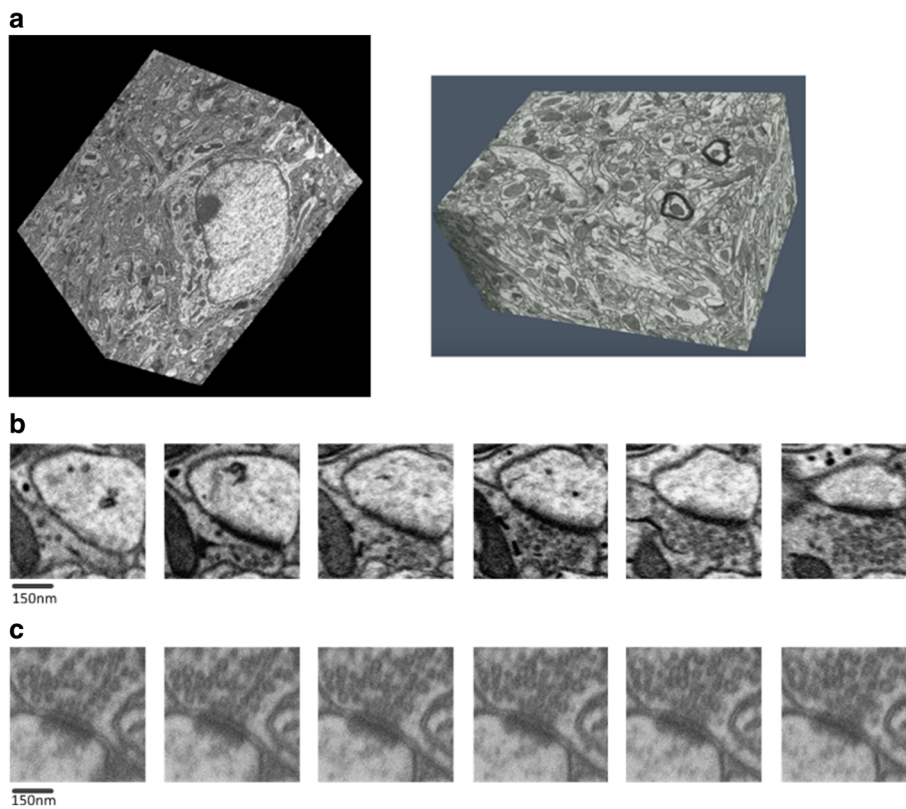


Fig. 1 Datasets and synapses. **a** Left: An anisotropic stack of neural tissue from mouse cortex acquired by ATUM-SEM. Right: Isotropic physical sections from rat hippocampus obtained by FIB-SEM. **b** Serial synapses in ATUM-SEM images. **c** Serial synapses in FIB-SEM images. As shown, the ATUM-SEM images are the sharper ones

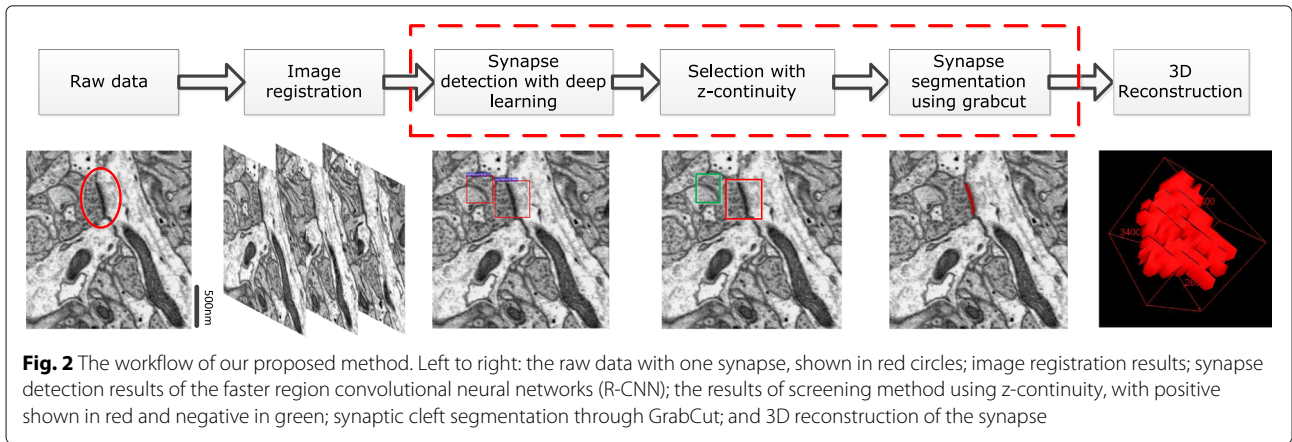
into five parts, as follows: image registration (ATUM-SEM only), synapse detection with deep learning, screening method with z-continuity, synapse segmentation using GrabCut and 3D reconstruction. The related video of the 3D reconstruction is shown in Additional file 1: Video S1. In this paper, we focus on the middle three steps. Figure 2 illustrates the workflow of the proposed method.

The proposed image registration method for serial sections of biological tissue is divided into three parts: searching for correspondences between adjacent section, displacement calculations for the identified correspondences, and warping the image tiles based on the new position of these correspondences. For the correspondences searching, we adopted SIFT-flow algorithm [31], to search for correspondences between adjacent sections by extracting equally distributed grid points on the well-aligned adjacent sections. For the displacement calculation, the positions of the identified correspondences were adjusted throughout all sections by minimizing a target energy function, which consisted of the data term, the small displacement term, and the smoothness term. The data term keeps pairs of correspondences at the same positions in the x-y plane after displacement. The small

displacement term constrains the correspondence displacements to minimize image deformation. The smoothness term constrains the displacement of the neighbor correspondences. For the image warping, we used the Moving Least Square (MLS) method [32] to warp each section with the obtained positions. The deformation results produced by MLS are globally smooth to retain the shape of biological specimens. The similar statement also can be seen from Ref. [33]. This image registration method not only reflects the discontinuity around wrinkle areas but also retains the smoothness in other regions, which provides a stable foundation for follow-up works.

Synapse detection with deep learning

In this part, Faster R-CNN was adopted to detect synapses in EM image stacks. Faster R-CNN mainly consists of two modules: the first module is the region proposal network (RPN), which generates region proposals, and the second one is Fast R-CNN [34], which classifies the region proposals into different categories. The process of applying Faster R-CNN to detect synapses is illustrated in Fig. 3. First, we used a shared fully convolutional network (FCN) to obtain the feature maps of the raw data. The



visualizations of feature maps indicate that, more neurons in the convolutional layer positively react to the visual patterns of synapses than others. Thus making it easier to recognize synapses from these maps. Subsequently, we adopted RPN to extract candidate regions from the feature maps (the architectures of shared FCN layers and RPN are illustrated in [Appendix 1]). Given the proposed regions and feature maps, the Fast R-CNN module was employed to classify the region proposals into synapse and background. In Faster R-CNN, the four basic steps of target detection, namely, region proposal, feature extraction, object classification and bounding-box regression, are unified in a deep-learning-based and end-to-end object detection system. Consequently, it is capable of

guaranteeing a satisfactory result in terms of both overall detection accuracy and operation speed.

Faster R-CNN is widely used to train and test natural image datasets, such as PASCAL VOC and MS COCO, where the height and width ranges of these images are from 500 pixels to 800 pixels. For an EM image, its size is generally larger than that of a natural image, larger than even 8000 pixels, which requires more memory storage in the GPU. To avoid exceeding the memory of the GPU, smaller images are proposed to train Faster R-CNN. For the ATUM-SEM dataset, we divided the original ATUM-SEM images (size of 8624 × 8416) into 72 small images (size of 1000 × 1000), allowing a nearly 50 pixel overlap between each image to avoid false negatives, as shown in

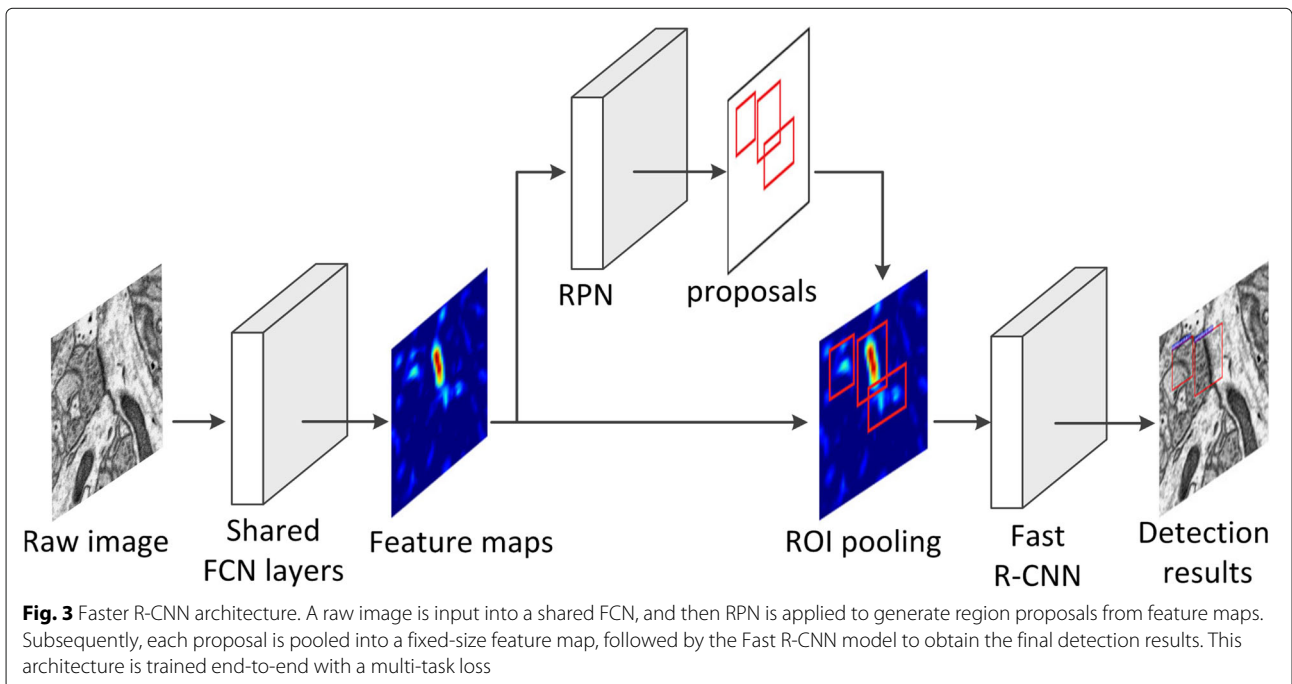
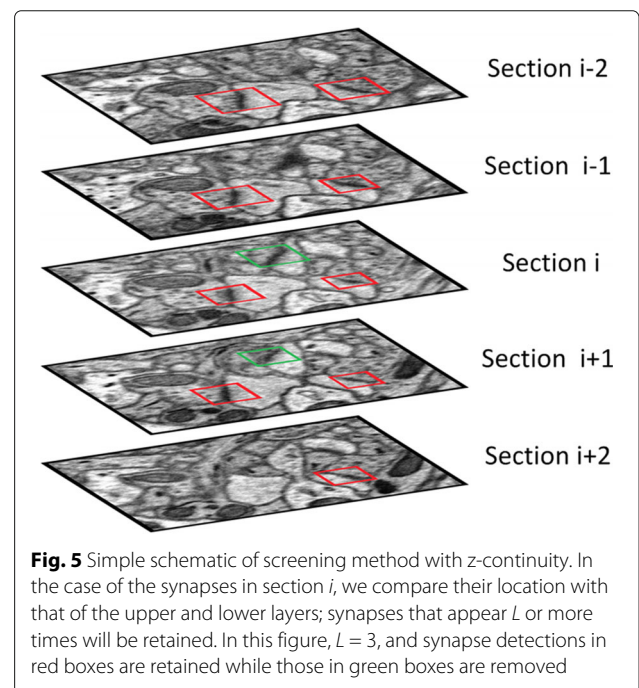
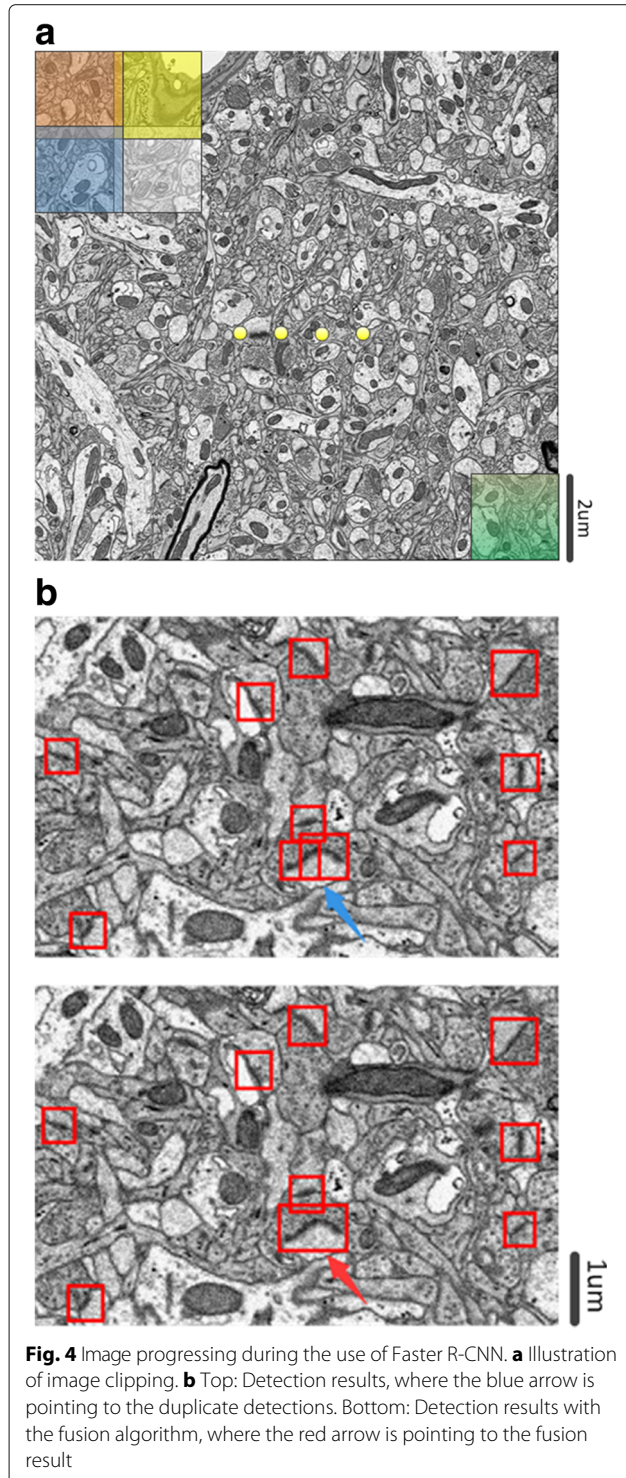


Fig. 4a. Similarly, we divided one original FIB-SEM image (size of 768×1024) into 6 overlapping small images (size of 500×500). In the following, the application Training Image Labeler was employed to label synapses. To avoid overfitting, we used augmentation strategy such as flip

and rotation to enlarge the training dataset. Through data augmentation, the number of both training samples is greater than 7000, which is sufficient for single target detection.

The deep learning network was implemented using Caffe [35] deep learning library (The process of training the Faster R-CNN is shown in [Appendix 2]). In training process, Faster R-CNN was optimized by the stochastic gradient descend (SGD) algorithm with the following optimization hyperparameters: weight decay = 0.0005, momentum = 0.9, gamma = 0.1, learning rate = 0.0001 for numerical stability. The mini-batch size and number of anchor locations were set to 128 and 2400, respectively. In addition to ZF [36] and VGG16 [37], we also applied ResNet50 [38] as shared FCN to train Faster R-CNN. It took nearly 20-28 hours to train the network for 80000 iterations on a GeForce Titan X GPU.

Given the detection results of small images, it is easy to gather all detections and obtain the final detection results of an original image. However, synapses distribute randomly in EM images, and it is possible that one synapse coexists in two adjacent small images. In this case, this method might lead to duplicate detections, which reduces the detection precision, as illustrated in Fig. 4b. Therefore, an effective detection boxes fusion method is proposed to solve this challenge. Through observations and analyses, we find that the distributions of synapses are sparse. Suppose that there are \mathcal{N}_i synapse detection boxes in the i th section, $\mathcal{S}_{i,j}$ represents the j th synapse detection box in the i th section, and $(c_{i,j}^1, c_{i,j}^2)$ and $(c_{i,j}^3, c_{i,j}^4)$ are



the upper-left coordinates and lower-right coordinates of $S_{i,j}$, respectively. If two synapse detection boxes are close enough or even overlapped, it can be concluded that these might be duplicate detections. A direct evaluation criterion for duplicate detections is the distance between synapses in the same section. The main procedure in the i th section is illustrated in Algorithm 1. In line 11 and 12 of Algorithm 1, $(c_{i,j}^1, c_{i,j}^2)$ and $(c_{i,j}^3, c_{i,j}^4)$ are the upper-left coordinates and lower-right coordinates of the updated $S'_{i,j}$, respectively.

Algorithm 1: Fusion of duplicate synapse detection boxes

Input:

\mathcal{N}_i : the number of synapse detection boxes in the i th section.
 $S_{i,j}, j \in [1, \mathcal{N}_i]$: j th synapse detection box in i th section.
 $C_{i,j}$: the coordinates of central point of $S_{i,j}$.
 ϑ : threshold.

Output:

$S'_{i,j}$: the updated j th synapse detection box in i th section.

```

1 Initialize  $j = 1$ 
2 repeat
3   for each  $S_{i,k} (j + 1 \leq k \leq \mathcal{N}_i)$  do
4     Calculate the Euclidean distance between  $S_{i,j}$ 
      and  $S_{i,k}$ :
5      $d_{i,j,k}^i = \|C_{i,j} - C_{i,k}\|^2, k = j + 1, \dots, \mathcal{N}_i.$ 
6   end
7   Seek the nearest synapse box  $S_{i,k_0}$  from  $S_{i,j}$ :
8    $k_0 = \operatorname{argmin} \{d_{i,j,k}^i\}, k = j + 1, \dots, \mathcal{N}_i.$ 
9   if  $d_{i,j,k_0}^i < \vartheta$  then
10    Fuse these two detection boxes  $S_{i,j}$  and  $S_{i,k_0}$ 
      into  $S'_{i,j}$ :
11     $c_{i,j}^r = \min(c_{i,j}^r, c_{i,k_0}^r), r = 1, 2$ 
12     $c_{i,j}^s = \max(c_{i,j}^s, c_{i,k_0}^s), s = 3, 4.$ 
13  end
14   $j \leftarrow j + 1$ 
15 until  $j = \mathcal{N}_i$ 

```

Screening method with z-continuity

A synapse is a flat 3D structure with a size of nearly 400 nm in long axis [39], whereas the distance between adjacent section is 50 nm in ATUM-SEM image stacks and 5 nm in FIB-SEM image stacks. As shown in Fig. 5, it can be hypothesized that a real synapse is capable of appearing in several layers.

In contrast, false positives only appear in one or two layers. Therefore, we utilized z-continuity to eliminate false positives. Specifically, if a synapse detection box appears L times or more in the same area of continuous $2L - 1$ layers, it can be considered as a real synapse; otherwise, it is regarded as a false positive. The clear-cut principle is described in Algorithm 2.

Algorithm 2: Screening method with z-continuity

Input:

\mathcal{M} : the number of all images.
 \mathcal{N}_i : the number of synapse detection boxes in the i th section.
 $S_{i,j}, j \in [1, \mathcal{N}_i]$: j th synapse detection box in i th section.
 $C_{i,j}$: the coordinates of central point of $S_{i,j}$.
 ν : distance threshold.
 L : z-continuity layers.

Output:

$S''_n (n = 1, 2, \dots)$: the screening results.

```

1 Initialize  $i = L + 1, j = 1$ 
2 repeat
3   repeat
4     for each
5        $S_{l,m} (i - L \leq l \leq i + L - 2, 1 \leq m \leq \mathcal{N}_l)$  do
6         Calculate the Euclidean distance between
           $S_{i,j}$  and other synapse detection boxes  $S_{l,m}$ 
          in continuous  $2L + 1$  layers:
7          $d_{j,m}^{i,l} = \|C_{i,j} - C_{l,m}\|^2, l =$ 
           $i - L, \dots, i + L; m = 1, \dots, \mathcal{N}_l.$ 
8       end
9       Determine the nearest synapse box
         $S_{l,t_l}, l = i - L, \dots, i + L.$  from  $S_{i,j}$  in each layer:
10       $t_l = \operatorname{argmin} \{d_{j,m}^{i,l}\}, m = 1, \dots, \mathcal{N}_l.$ 
11      Obtain the times that the synapse detection
        box  $S_{i,j}$  appears:
12       $T = \sum_{l=i-L}^{i+L} [d_{j,t_l}^{i,l} < \nu].$ 
13      if  $T \geq L$  then
14        Remain the synapse detection box  $S_{i,j}$ .
15      end
16       $j \leftarrow j + 1$ 
17    until  $j = \mathcal{N}_i$ 
18  until  $i = \mathcal{M} - L;$ 

```

In line 11 of Algorithm 2, $[\cdot]$ denotes the indicator function. When T meets $T \geq L$, we can confirm that the object detected by Faster R-CNN is a synapse with high probability; thus, this detection result $S_{i,j}$ with an index in the 3D view $S''_n (n = 1, 2, \dots)$ will remain. Otherwise, it will be removed.

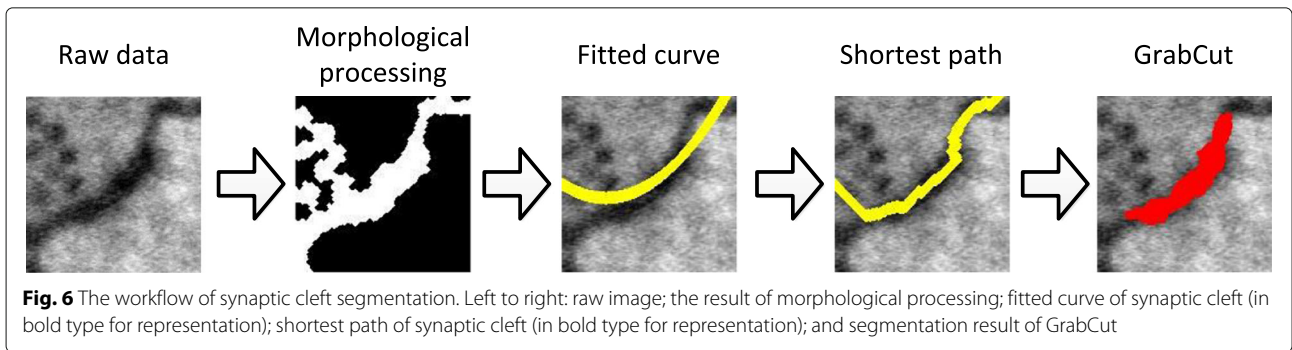


Fig. 6 The workflow of synaptic cleft segmentation. Left to right: raw image; the result of morphological processing; fitted curve of synaptic cleft (in bold type for representation); shortest path of synaptic cleft (in bold type for representation); and segmentation result of GrabCut

Synapse segmentation using GrabCut

Because synaptic clefts, which are at least 40 nm in width, are wider than other dark regions in the detection boxes, they can be segmented using several image processing methods, as illustrated in Fig. 6.

First, we converted the original detection images into binary images using an adaptive threshold. On this basis, the erode and dilate operations were employed to eliminate noise and obtain synaptic clefts. After morphological processing, synaptic clefts can be approximately located. Since most shapes of synaptic clefts are similar to quadratic curves, suitable curves are proposed to fit the structure of the synaptic clefts and obtain more refined results. We randomly selected m pairs of points $p_i = (x_i, y_i), 1 < i \leq m$ from the image after morphological processing, and m is defined as one third of the number of white points in the corresponding image, which is empirically based. Subsequently, we employed them to fit the quadratic curve $y = ax^2 + bx + c$.

Consequently, a series of synaptic clefts are observed as quadratic curves. Finally, we selected the starting point p_1 and the ending point p_n from the two ends of each fitted curve, and then we calculated the shortest path [28] between p_1 and p_n .

Note that the obtained shortest path is only a curve rather than a segmentation result, and sometimes the dilated results of fitted curve and shortest path cannot effectively fit the various synaptic clefts, as shown in the Fig. 6, an effective segmentation algorithm has to be introduced. Motivated by previous researches [29], we proposed to use GrabCut algorithm for fine segmentation.

First, we considered the image as an array $\alpha = (\alpha_1, \dots, \alpha_N)$, and we regarded the segmentation result as an array $\beta = (\beta_1, \dots, \beta_n, \dots, \beta_N)$ at each pixel, $\beta_n = \{0, 1\}$ with 1 for synapse and 0 for background. The parameters θ denoted the distributions of synapse and background in the image. Next, we modeled a full-covariance Gaussian mixture model (GMM) [40] for synapse and background

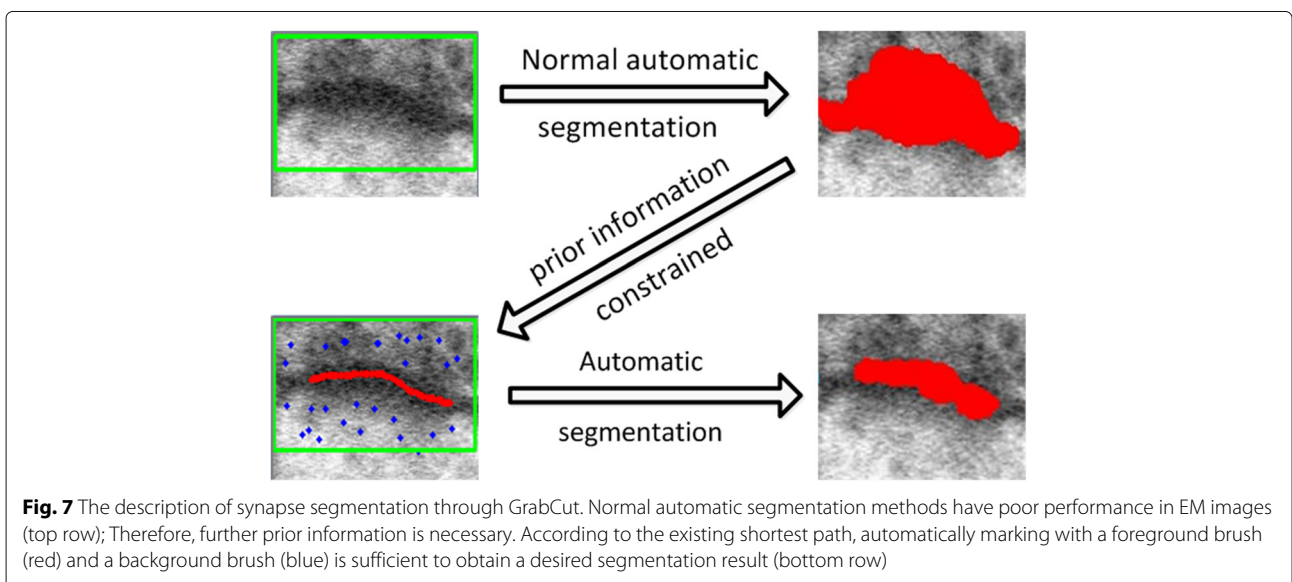


Fig. 7 The description of synapse segmentation through GrabCut. Normal automatic segmentation methods have poor performance in EM images (top row); Therefore, further prior information is necessary. According to the existing shortest path, automatically marking with a foreground brush (red) and a background brush (blue) is sufficient to obtain a desired segmentation result (bottom row)

with K components separately. To properly use the GMM, an additional vector $k = (k_1, \dots, k_n, \dots, k_N)$ is introduced, where k_n is the Gaussian component corresponding to the n th pixel. For each pixel, the GMM component is either from the synapse model or the background model. The task of segmentation is to obtain the unknown variables $\underline{\beta}$ from the given image α and the model parameters $\underline{\theta}$. The Gibbs energy E consists of a data term \mathcal{D} and smoothness term \mathcal{S} , which can be defined as

$$E(\underline{\beta}, k, \underline{\theta}, \alpha) = \mathcal{D}(\underline{\beta}, k, \underline{\theta}, \alpha) + \mathcal{S}(\underline{\beta}, \alpha). \quad (1)$$

The minimum of Eq. (1) can be considered as a good segmentation. In Eq. (1), the data term \mathcal{D} indicates the penalty for a pixel that is classified incorrectly. According to the GMMs, it can be defined as

$$\mathcal{D}(\underline{\beta}, k, \underline{\theta}, \alpha) = \sum_n G(\underline{\beta}_n, k_n, \underline{\theta}, \alpha_n). \quad (2)$$

where G can be expressed as

$$G(\underline{\beta}_n, k_n, \underline{\theta}, \alpha_n) = -\log p(\alpha_n | \underline{\beta}_n, k_n, \underline{\theta}) - \log \pi(\underline{\beta}_n, k_n). \quad (3)$$

In this work, $p(\cdot)$ denotes the Gaussian probability distribution, and $\pi(\cdot)$ represents the mixture weight.

GraphCut [41] is a one-time minimization, whereas GrabCut is an iterative minimization, and each iteration process makes the GMM parameters better for image segmentation. Initialize the trimap $T = \{T_S, T_B, T_U\}$ by selecting the rectangular box. The pixels outside the box belong to background T_B , whereas the pixels inside the box indicate “they might be synapses” and belong to T_U , and T_S implies synapse. To obtain a better result, users can draw a masking area inside the box with a synapse brush and a background brush, where the pixels in different masking areas are regarded as different classes. The detailed procedures for synapse segmentation using GrabCut are described in Algorithm 3.

In Fig. 7, we present a more visual description of synapse segmentation by using GrabCut. The bounding box in green is automatically obtained by the boundary of images, which denotes the initial area T_U (top left corner). In the case of our datasets, the automatic segmentation result is not accurate (top right corner). For this task, we take the skeletonization of the shortest path and random sampling points as prior information, and then we apply GrabCut to synapse segmentation. The skeletonizations of the shortest path (in red) are regarded as synapse T_S (bottom left corner), and random sampling points (in blue) indicate background T_B . In this case, the final segmentation result is satisfactory.

Algorithm 3: Iterative image segmentation with GrabCut

Input:
 $\alpha = (\alpha_1, \dots, \alpha_N)$: image data.
 $T = \{T_S, T_B, T_U\}$: the initialized trimap.
Output:
 $\beta = (\beta_1, \dots, \beta_n, \dots, \beta_N)$: segmentation result.

- 1 **Initialize** the k-means algorithm is adopted to initialize the synapse and the background GMM.
- 2 **repeat**
- 3 Assign GMM components to each pixel in T :
- 4 $k_n := \arg \min_{k_n} G_n(\beta_n, k_n, \theta, \alpha_n)$.
- 5 Learn and optimize GMM parameters through image data α :
- 6 $\underline{\theta} := \arg \min_{\underline{\theta}} \mathcal{D}(\underline{\beta}, k, \underline{\theta}, \alpha)$.
- 7 Utilize min cut to minimize energy E and obtain the segmentation result:
- 8 $\min_{\{\beta_n: n \in T_U\}} \min_k E(\underline{\beta}, k, \underline{\theta}, \alpha)$
- 9 **until convergence**;

Results and Discussion

In this section, we present several experiments on the two datasets depicted in Fig. 1a to validate our proposed method. All algorithm parameters are summarized in Table 1. Due to the difference between the two datasets, the parameters are different in the detection box fusion process and screening method with z-continuity. Parameters (subscript 1) are applied to the ATUM-SEM dataset, and parameters (subscript 2) are suitable for the FIB-SEM dataset. Although it appears that there are many parameters to tune in the pipeline, only the fusion distance threshold, z-continuity layer and z-continuity distance threshold need to be replaced when the dataset changes.

We first present the datasets and evaluation methods. Subsequently, we adopt precision-recall curves to evaluate the performances of our detection and segmentation

Table 1 Algorithm parameters and values

Parameter	Symbol	Value
Fusion distance threshold	ϑ_1	100
	ϑ_2	50
z-continuity layers	L_1	3
	L_2	20
z-continuity distance threshold	v_1	200
	v_2	100
GrabCut iterations	T_G	10
Gaussian mixture components	K	5

Table 2 Illustration of two datasets

Dataset	EM	Voxel size (nm^3)	Train size	Test size
(A) Cortex	ATUM-SEM	$2 \times 2 \times 50$	$8624 \times 8416 \times 30$	$7616 \times 8576 \times 178$
(B) Hippocampus	FIB-SEM	$5 \times 5 \times 5$	$1024 \times 768 \times 165$	$1024 \times 768 \times 165$

methods. Then, average precision (AP), F1 score and Jaccard index are employed for further quantitative analyses. Finally, we present and analyze the reconstruction results of synapses.

Datasets and evaluation method

In this work, the specimens and ATUM-SEM sections of mouse cortex were provided by the Institute of Neuroscience, Chinese Academy of Sciences (CAS). The physical sections were imaged using an SEM (Zeiss Supra55) with an imaging voxel resolution size of $2 \text{ nm} \times 2 \text{ nm} \times 50 \text{ nm}$ and dwell time of $2 \mu\text{s}$ by the Institute of Automation, CAS. The dataset of rat hippocampus¹ was acquired by Graham Knott and Marco Cantoni at École Polytechnique Fédérale de Lausanne (EPFL). It is made publicly available for accelerating neuroscience research, and the resolution of each voxel is approximately $5 \text{ nm} \times 5 \text{ nm} \times 5 \text{ nm}$.

The details of the training and testing data for each dataset are summarized in Table 2. The ground truths of

the datasets are annotated manually using ImageJ software. For the ATUM-SEM dataset, the training dataset contains 142 synapses in 3D view and 1522 synapses in 2D view, and the testing volume contains 723 synapses in 3D view and 7183 synapses in 2D view. For the FIB-SEM dataset, the number of synapses in training and testing are 25 and 26. Clearly, it is a time consuming and laborious process, which takes several experienced students approximately one month to obtain such a large amount of databases, respectively.

Since our approach is composed of two primary parts, detection and segmentation, we choose different metrics for the different parts. The main metrics utilized for evaluation are as follows:

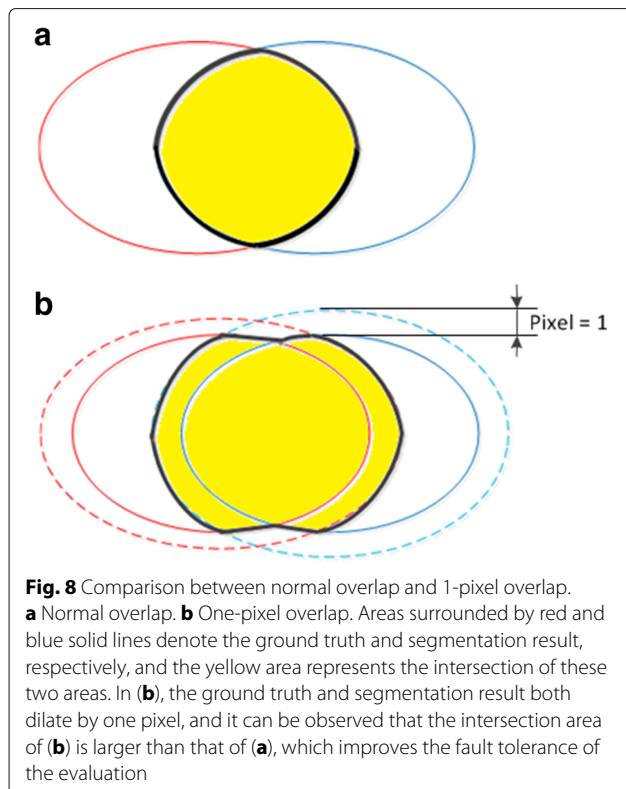
- **Precision and recall.** In this work, precision is the probability that detected synapses are correct, and recall is the probability that the true synapses are successfully detected.

$$\text{precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}, \quad (4)$$

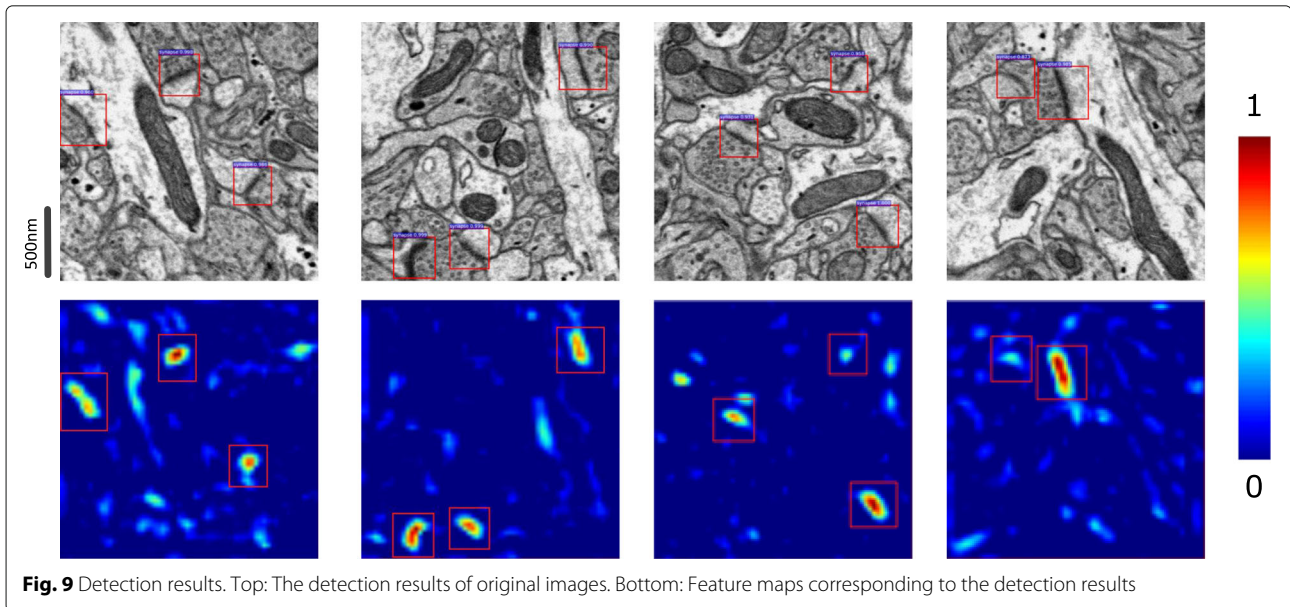
$$\text{recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}. \quad (5)$$

- **Average precision.** AP denotes the area under the precision-recall curve, and it can be expressed as the following formula, where P represents precision and R indicates recall:

$$AP = \int_0^1 P(R) dR. \quad (6)$$

**Table 3** Detection results of Faster R-CNN based on different models

Dataset	EM	Size	Model	AP	Rate
(A) Cortex	ATUM-SEM	1000×1000	ZF	82.0%	9 fps
			VGG16	83.2%	3 fps
			ResNet50	84.1%	4 fps
(B) Hippocampus	FIB-SEM	500×500	ZF	86.8%	36 fps
			VGG16	87.4%	12 fps
			ResNet50	90.9%	16 fps



- *F1 score*. Since precision and recall are often contradictory, *F1 score* is the weighted average of precision and recall, which shows the comprehensive performance of methods.

$$F1\ score = \frac{2 \times P \times R}{P + R}. \tag{7}$$

- *Jaccard index*. This metric is also known as the VOC score [42], which calculates the pixel-wise overlap between the ground truth (Y) and segmentation results (X).

$$Jaccard\ index(X, Y) = \frac{X \cap Y}{X \cup Y}. \tag{8}$$

Motivated by Ref. [43], we define that a detection or segmentation result is considered as a true positive only if the overlap between the region of the result and corresponding ground truth reaches at least 70%.

For segmentation, the shape of the synapses is always long and narrow, and the boundaries of synapses are often difficult to define. According to Ref. [15], manual annotations near synapse borders are not always accurate. Hence, due to the error in the annotations, the evaluation measure such as Jaccard index may be impacted with high probability. Inspired by the average 3-pixel error rate in Ref. [44], we define a pixel neighborhood overlap measure to eliminate this adverse effect. As depicted in Fig. 8a, the area surrounded by the red solid line denotes the ground truth (Y), and the area surrounded by the blue solid line indicates the segmentation result (X). The yellow area in Fig. 8a represents the intersection of the ground truth

and segmentation result. In Fig. 8b, both the ground truth and segmentation result dilate one pixel, and the dilated ground truth (Y^1) and dilated segmentation result (X^1) are denoted with dashed lines. Therefore, the Jaccard index of 1-pixel overlap can be expressed as

$$Jaccard\ index^1(X, Y) = \frac{(X^1 \cap Y^1) \cup (X \cap Y^1)}{X \cup Y}. \tag{9}$$

Furthermore, we use the dilated segmentation result (X^1) and dilated ground truth (Y^1) to calculate the precision and recall and to obtain the 1-pixel overlap of AP. For 3-pixel overlap and 5-pixel overlap, the ground truth and segmentation result dilate three pixels and five pixels, respectively.

Detection Accuracy

In this subsection, we evaluate the detection performance of our approach and compare it with Refs. [15, 18, 23] on different datasets in terms of precision recall curves, AP and F1 measure.

Table 3 presents the detection results of Faster R-CNN on a GeForce Titan X GPU. In Table 3, the rate shows the processing speed of different models on the test images. Through the experimental results, it can be found that the ResNet50 network provides the highest AP with an acceptance rate. Therefore, we exploited ResNet50 networks to detect synapses.

Table 4 Detection performance of different threshold L on the ATUM-SEM dataset

Size of L	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$L = 6$	$L = 7$
AP	77.8	86.8	83.3	75.2	70.1	66.2

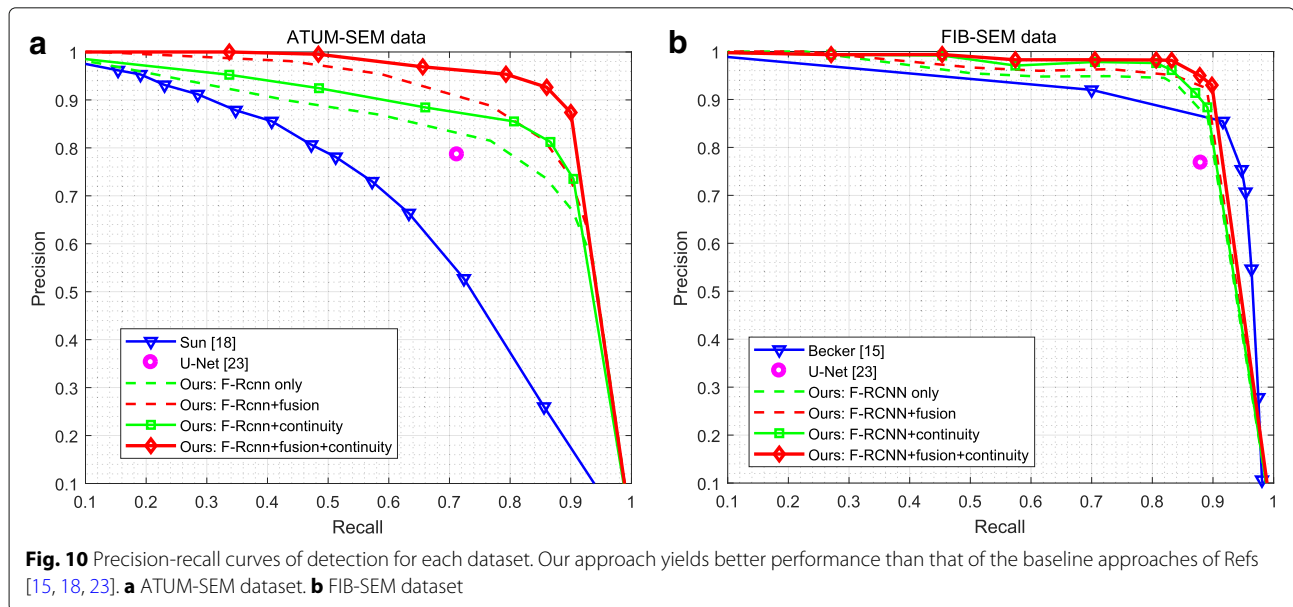


Figure 9 shows the detection results of original images and presents the feature maps extracted from the Faster R-CNN model. These figures indicate that neurons in the convolutional layer positively react to the visual patterns of synapses.

From the subsection of screening method with z-continuity, it can be hypothesized that a real synapse is capable of appearing L times or more in the same area of continuous $2L - 1$ layers. Therefore, we conduct several experiments by considering the cases of $L = 2, 3, 4, 5, 6, 7$ on the ATUM-SEM dataset. As shown in Table 4, the detection performance (measured by AP) is highest when taking $L = 3$. When $L \geq 3$, note that the AP decrease as the threshold L increase, which is mainly because of filtering the true positives.

Precision-recall curves are presented in Fig. 10. According to the behavior of the precision-recall plots, it can be deduced that our approach performs better than the baseline approaches of Refs. [15, 18, 23] for all recall values.

For the ATUM-SEM dataset, the performance significantly decreases since it ignores fusion and z-continuity. For the FIB-SEM dataset, the difference between synapses and other subcellular structures is significant due to its small area and simple scene. Thus, our approach achieves a higher average detection precision, and the promotion of the screening method is not enormous.

Although U-Net [23] is a simple and effective fully convolutional networks for image segmentation, which achieves good performance on different biomedical segmentation applications. Due to the imbalanced data

Table 5 Quantitative detection performance for two EM datasets

Metrics	Sun [18]	U-Net [23]	F-RCNN	F-RCNN + fusion	F-RCNN + continuity	F-RCNN + fusion + continuity
ATUM-SEM DATASET						
AP	67.6%	74.9%	84.1%	89.7%	86.8%	92.8%
F1	64.8%	74.8%	79.3%	83.6%	83.9%	89.2%
FIB-SEM DATASET						
Metrics	Becker [15]	U-Net [23]	F-RCNN	F-RCNN + fusion	F-RCNN + continuity	F-RCNN + fusion + continuity
AP	90.6%	82.4%	90.9%	91.9%	92.4%	93.5%
F1	88.4%	82.0%	88.5%	90.6%	89.2%	91.4%

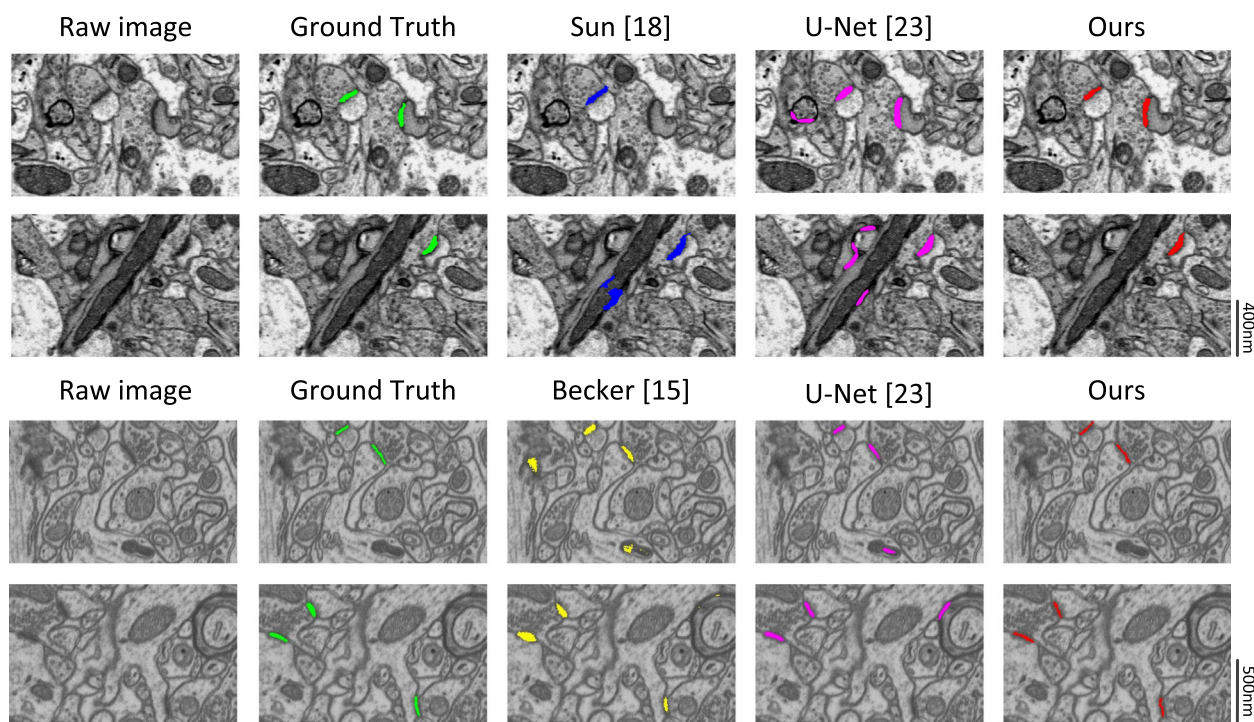


Fig. 11 Qualitative segmentation results for two different datasets at the best threshold. Top: ATUM-SEM images, Bottom: FIB-SEM images. Note that our method achieves more accurate results as well as reducing false positives

and complex scene, the performance of U-Net is not very satisfactory in synapse segmentation task. As shown in Table 5, our approach outperforms other algorithms on different metrics. For the ATUM-SEM and FIB-SEM datasets, our approach yields 92.8 and 93.5% AP and 89.2 and 91.4% F1 score, which are higher than that of others algorithms.

Segmentation accuracy

The qualitative segmentation results are shown in Fig. 11, from which we can easily conclude that our approach achieves satisfactory results. Our approach provides more accurate results and reduces false positives, and most of the synaptic clefts are correctly segmented from the original images.

For the quantitative analysis, evaluation of the model is performed by using precision-recall curves, AP, F1 score and Jaccard index. Precision-recall curves at different pixel overlap sizes are shown in Fig. 12, and our approach outperforms the baseline approaches of Refs. [15, 18, 19, 23] for most of the recall values and pixel overlap sizes. Note that the proposed approach with GrabCut perform much better than the approach with variational region growing [19], which verifies the effectiveness of GrabCut. In the following, we present the AP and the highest values of the F1 score and Jaccard index for all recalls at

different values of pixel overlap sizes in Tables 6 and 7, which illustrate the same results. In general, our results indicate a significant improvement in the segmentation performance for each dataset.

3D visualization

After obtaining the segmentation results, we import them into ImageJ [30], and we show the 3D visualization of synapses for each dataset in Fig. 13. The related videos are shown in Additional file 2: Video S2 and Additional file 3: Video S3, respectively.

Computational Efficiency

In this subsection, we illustrated the comparisons of computational efficiency in Table 8. Being different from Ref. [18] and our approach, Ref. [15] obtained the synapse segmentation results directly, thus its computational cost only contain segmentation part. For the large-scale ATUM-SEM dataset, the total detection and segmentation time of our approach is 255 mins, which is slightly slower than that of Ref. [18] (202 mins). For the FIB-SEM dataset, the test time of our approach is 8 mins, which is almost one-fourth of Ref. [15]. Note that our approach outperforms Refs. [15, 18] in precision of detection and segmentation. It can be speculated that the proposed pipeline is capable of guaranteeing a promising result

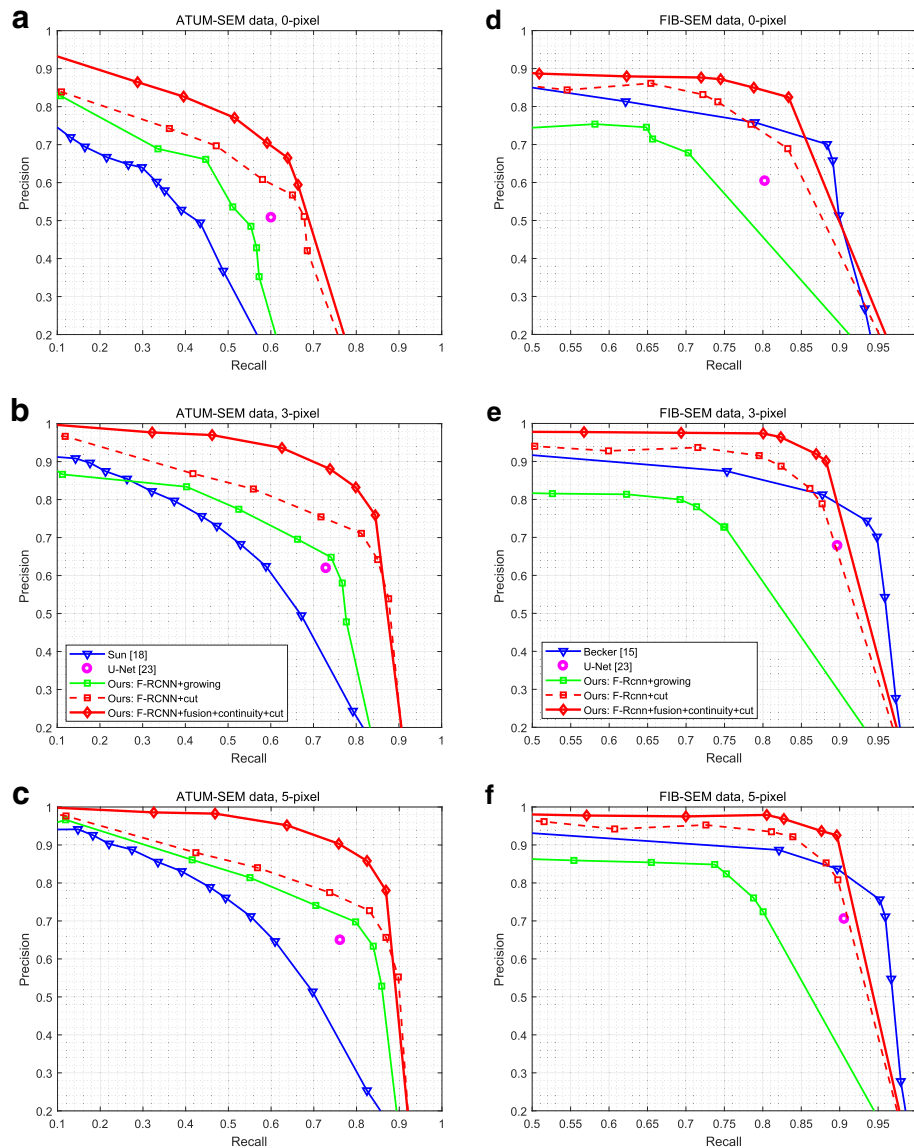


Fig. 12 Precision-recall curves of segmentation for each dataset at different pixel overlap sizes. Cut denotes the GrabCut method, growing denotes the variational region growing method. The performance of our approach is better than that of the baseline approaches of Refs. [15, 18, 19, 23]. **a-c** ATUM-SEM dataset at 0, 3 and 5 pixel overlap, respectively. **d-f** FIB-SEM dataset at 0, 3 and 5 pixel overlap, respectively

in terms of both segmentation accuracy and operation speed.

Discussion

As mentioned in this section, our approach outperforms the approaches of Refs. [15, 18, 19, 23] on several standard metrics. However, note that the results of Ref. [15] in this paper are lower than those reported in the TMI paper. There might be two reasons for this inaccuracy. The first is that the authors of Ref. [15] offer no ground truth, and they allow us to draw it by ourselves. The second is that

our performance measurements are not similar to those in Ref. [15].

Since the synapse is a flat 3D structure and the screening method with z-continuity have indeed reduced false positives in our work, which demonstrates the importance of 3D information in synapse detection problem. In addition, inspired by the promising results [45, 46], it can be speculated that the 3D network could effectively preserve and extract the 3D spatial information from volumetric data. Therefore, we believe that the extension of 2D R-CNN to 3D one could help improve the detection

Table 6 Segmentation performance for the ATUM-SEM dataset

Size of pixel overlap	Metrics	Sun [18]	U-Net [23]	F-RCNN+ Growing [19]	F-RCNN + GrabCut	F-RCNN + fusion + continuity + GrabCut
0-pixel	AP	32.7%	55.4%	49.0%	57.6%	65.6%
	F1	36.3%	58.1%	53.4%	60.6%	65.2%
	Jaccard	30.6%	49.5%	43.2%	52.6%	58.1%
3-pixel	AP	56.3%	67.4%	67.8%	78.8%	86.1%
	F1	60.6%	66.9%	69.2%	75.8%	81.5%
	Jaccard	50.1%	59.0%	59.8%	69.0%	74.7%
5-pixel	AP	59.4%	70.6%	76.4%	80.3%	88.6%
	F1	62.7%	70.1%	74.4%	77.5%	84.1%
	Jaccard	53.7%	60.2%	63.5%	72.2%	76.9%

accuracy, and we are planning to design 3D Faster R-CNN network to detect synapses in EM data set.

Conclusion

In this paper, we propose an effective approach to reconstruct synapses with deep learning. Our strategy is to utilize Faster R-CNN to detect the regions of synapses and then employ z-continuity to reduce false positives. Subsequently, shortest path and GrabCut are employed to obtain the synaptic clefts. Finally, we utilize our approach for the 3D reconstruction of synapses in isotropic and anisotropic datasets. The experimental results demonstrate that our algorithm enhances the precision of detection and guarantees the accuracy of segmentation, which will promote efficiency in synapse validation and benefit connectomics and synaptic plasticity analysis. Furthermore, we apply our approach to neuroscience experiments. Our automated approach helps neurologists quickly identify the number of synapses and multi-synapses in different experimental specimens, and further

analyses reveal a correlation between spine formation and responses of fear-conditioned animals [4].

Despite the promising segmentation results of our approach, the segmentation process is somewhat tedious, and the efficiency and accuracy of traditional segmentation algorithms can be increased. For this task, future work will focus on detection and segmentation using end-to-end 3D deep neural networks, which will enhance both speed and accuracy for synapse reconstruction algorithm.

Endnote

¹ <http://cvlab.epfl.ch/data/em>

Appendix 1: The architecture of shared FCN layers and RPN

Faster R-CNN mainly consists of two modules, the first module is RPN that generates region proposals, the second one is Fast R-CNN [34] which classifies the region proposals into different categories. In practice, the architecture of RPN and Fast R-CNN are fixed, ZF [36] and

Table 7 Segmentation performance for the FIB-SEM dataset

Size of pixel overlap	Metrics	Becker [15]	U-Net [23]	F-RCNN + Growing [19]	F-RCNN + GrabCut	F-RCNN + fusion + continuity + GrabCut
0-pixel	AP	79.3%	70.3%	66.8%	79.4%	82.8%
	F1	78.2%	68.9%	69.4%	77.5%	82.9%
	Jaccard	61.5%	54.1%	53.5%	65.2%	70.9%
3-pixel	AP	88.0%	78.8%	74.0%	88.1%	91.6%
	F1	84.4%	77.2%	74.6%	85.4%	89.4%
	Jaccard	74.6%	60.9%	60.2%	75.7%	79.6%
5-pixel	AP	90.2%	80.6%	79.5%	90.8%	93.0%
	F1	86.6%	79.3%	78.9%	87.8%	91.1%
	Jaccard	80.3%	67.5%	67.7%	80.6%	83.7%

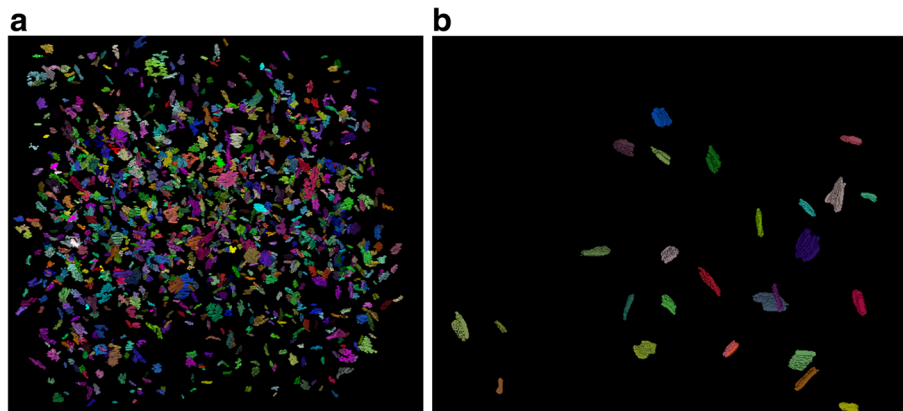


Fig. 13 3D reconstruction of synapses. **a** Reconstruction of synapses on the ATUM-SEM dataset with a volume of $15.2 \mu\text{m} \times 17.2 \mu\text{m} \times 8.9 \mu\text{m}$. **b** Reconstruction of synapses on the FIB-SEM dataset with a volume of $5.1 \mu\text{m} \times 3.8 \mu\text{m} \times 0.8 \mu\text{m}$

VGG16 [37] are applied as shared FCN to train the whole network. Here we introduce the shared FCN and RPN modules in details.

The architecture of the shared FCN layers (eg. ZF) and RPN is showed in Fig. 14. The structure above the dotted line is ZF network, and the structure below the dotted line is RPN network. Assuming the size of the input map is $224 \times 224 \times 3$. The parameters of Conv1 layer are as follows: the kernel size is 7×7 , the channels of feature maps is 96, the padding size and the stride size is 5 and 2, respectively. In ZF, each convolution layer is followed by batch normalization and rectified linear unit (RELU), and the size of output feature maps is $14 \times 14 \times 256$. In RPN, it first takes a 3×3 convolutional kernel (sliding window) in the input feature maps. In the center of sliding window (illustrated in Fig. 15), there are k ($k = 9$) anchor boxes with different scales (128, 256, 512) and aspect ratios (1:1, 1:2, 2:1), which are used to solve the multi-scale target tasks in detection. In what follows, *cls* layer and *reg* layer are employed for classification and border regression respectively. *cls* layer generates two elements to determine the probability of candidates, while the *reg* layer produces four coordinate elements (x, y, w, h) to identify the location of candidates. Finally, according to the probability of candidates, RPN selects the top 300

region proposals as the input of Fast R-CNN for further classification.

Appendix 2: The process of training the Faster R-CNN

Since RPN and Fast R-CNN share the same feature extraction network, we utilized the alternating training method to train the network with features shared, the training process is described as follows:

Step 1: Initialize RPN network parameters by using ImageNet pre-trained model, and then fine-tune RPN network.

Step 2: Utilize the region proposals generated by the step-1 RPN to train Fast R-CNN detection network. In this process, ImageNet pre-training model is also used to initialize Fast R-CNN network parameters.

Step 3: Keep the shared FCN layers to be fixed. Subsequently, apply step-2 Fast R-CNN network and train data to re-initialize RPN, and only fine-tune the layers unique to RPN.

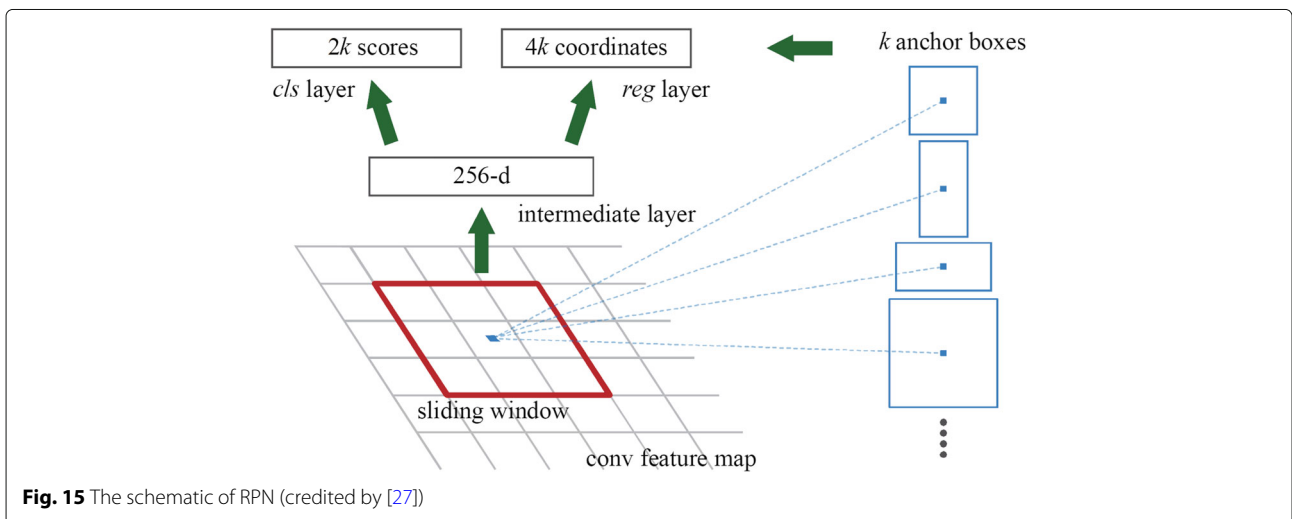
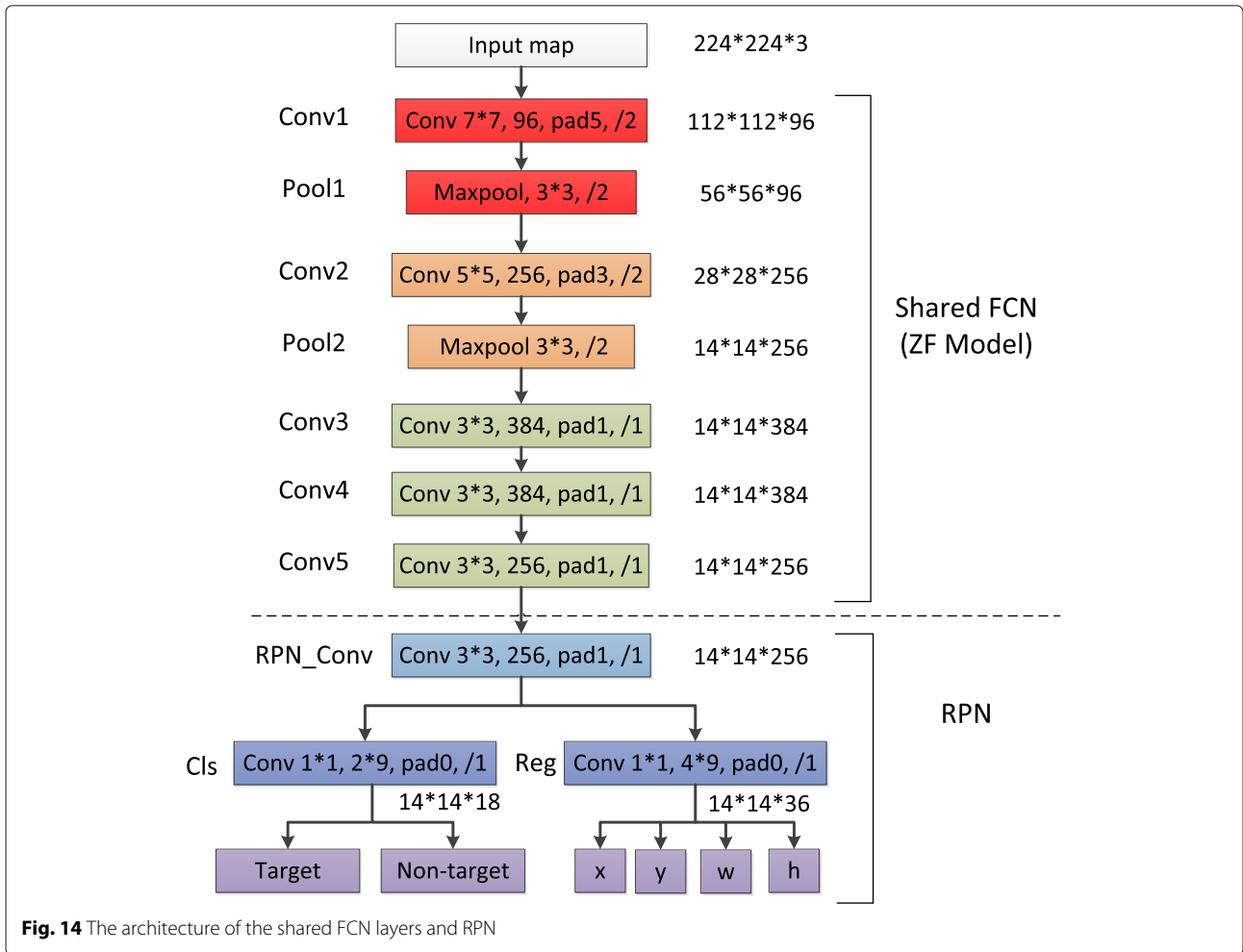
Step 4: Employ step-3 RPN and train data to fine-tune Fast R-CNN with the shared FCN layers fixed.

The above four steps enable RPN and Fast R-CNN networks to share the convolutional layers and form a unified Faster R-CNN network.

Table 8 The comparisons of computational efficiency

Dataset	Volume (μm^3)	Becker [15]		Sun [18]		Our approach	
		Segmentation		Detection	Segmentation	Detection	Segmentation
ATUM	$15.2 \times 17.2 \times 8.9$	-	-	48 mins	154 mins	70 mins	185 mins
FIB	$5.1 \times 3.8 \times 0.8$	30 mins	-	-	-	2 mins	6 mins

The test time was calculated on the same desktop equipped with an Intel i7-4790 CPU of 32 GB memory and a GeForce Titan X GPU



Additional files

Additional file 1: Video S1. 3D reconstruction result in Fig. 2. (MOV 946 kb)

Additional file 2: Video S2. 3D reconstruction of synapses on the ATUM-SEM dataset. (MOV 13,499 kb)

Additional file 3: Video S3. 3D reconstruction of synapses on the FIB-SEM dataset. (MOV 2275 kb)

Abbreviations

ATUM-SEM: Automated tape-collecting ultramicrotome scanning electron microscopy; AP: Average precision; CAS: Chinese Academy of Sciences; DNNs: Deep neural networks; EM: electron microscopy; EPFL: École Polytechnique Fédérale de Lausanne; FCN: Fully convolutional network; FIB-SEM: Focused ion beam scanning electron microscopy; GMM: Gaussian mixture model; MLS: Moving least squares; MSER: Maximally stable extremal region; R-CNN: Region convolutional neural networks; RPN: Region proposal network; SBEM: Serial block-face scanning electron microscopy; SIFT: Scale-invariant feature transform; ssTEM: Serial section transmission electron microscopy; SVM: Support vector machine; TMI: Transactions on medical imaging

Acknowledgements

We would like to thank the editors and reviewers for their valuable comments and suggestions that helped improve the original version of this paper, thank Dr. Yu Kong, Yang Yang and Danqian Liu (Institute of Neuroscience, CAS) for sample preparation and sectioning. We would also like to thank Mr. Lixin Wei and colleagues (Institute of Automation, CAS) for use of the Zeiss Supra55 Scanning Electron Microscope and technical support.

Funding

Data collection received support from Scientific Instrument Developing Project of Chinese Academy of Sciences (No. YZ201671). Analysis and interpretation received support from National Science Foundation of China (No. 61201050, No. 11771130). Design and writing of the manuscript received support from Special Program of Beijing Municipal Science and Technology Commission (No. Z16110000216146).

Availability of data and materials

The data and source code in this paper is available upon request.

Authors' contributions

CX conceived and designed the method, developed the algorithm and wrote the paper. WL, HD implemented the screening algorithm, conducted the experiments and wrote the paper. XC developed the image registration method. HH and QX conceived the method and gave the research direction and feedback on experiments and results. YY provided mouse cortex dataset and was instrumental in analysing the results. All authors contributed to, read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹Institute of Automation, Chinese Academy of Sciences, 95 Zhongguancun East Road, 100190 Beijing, China. ²School of Future Technology, University of Chinese Academy of Sciences, 19 Yuquan Road, Shijingshan District, 100049 Beijing, China. ³Faculty of Mathematics and Statistics, Hubei University, Institute of Automation, Chinese Academy of Sciences, 368 Youyi Road, Wuhan, China. ⁴Faculty of Information Technology, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau, China. ⁵Institute of

Neuroscience, Chinese Academy of Sciences, 320 Yue Yang Road, 200031 Shanghai, China. ⁶Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, 320 Yue Yang Road, 200031 Shanghai, China. ⁷Data Mining Lab, Beijing University of Technology, 100 Ping Le Yuan, 100124 Beijing, China.

Received: 16 December 2017 Accepted: 4 June 2018

Published online: 13 July 2018

References

- Briggman KL, Helmstaedter M, Denk W. Wiring specificity in the direction-selectivity circuit of the retina. *Nature*. 2011;471(7337):183–8.
- Helmstaedter M, Briggman KL, Turaga SC, Jain V, Seung HS, Denk W. Corrigendum: Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*. 2013;500(7461):168–74.
- Kasthuri N, Narayanan KJ, Hayworth DR, et al. Saturated reconstruction of a volume of neocortex. *Cell*. 2015;162(3):648–61.
- Yang Y, Liu DQ, Huang W, Deng J, Sun Y, Zuo Y, Poo MM. Selective synaptic remodeling of amygdalocortical connections associated with fear memory. *Nat Neurosci*. 2016;19(10):1348–55.
- Hofer SB, Mrsic-Flogel TD, Bonhoeffer T, Hübener M. Experience leaves a lasting structural trace in cortical circuits. *Nature*. 2009;457(7227):313–7.
- Zuo Y, Yang G, Kwon E, Gan WB. Long-term sensory deprivation prevents dendritic spine loss in primary somatosensory cortex. *Nature*. 2005;436(7048):261–5.
- Knott G, Marchman H, Wall D, Lich B. Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *J Neurosci*. 2008;28(12):2959–64.
- Hayworth KJ, Kasthuri N, Schalek R. Automating the collection of ultrathin serial sections for large volume TEM reconstructions. *Microsc Microanal*. 2006;12(S02):86–7.
- Lee PC, Chuang CC, Chiang AS, Ching YT. High-throughput computer method for 3d neuronal structure reconstruction from the image stack of the Drosophila brain and its applications. *PLoS Comput Biol*. 2012;8(9):e1002658.
- Mishchenko Y, Hu T, Spacek J, Mendenhall J, Harris KM, Chklovskii ADB. Ultrastructural analysis of hippocampal neuropil from the connectomics perspective. *Neuron*. 2010;67(6):1009–20.
- Denk W, Horstmann H. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol*. 2004;2(11):1900–9.
- Navlakha S, Suhan J, Barth AL, Barjoseph Z. A high-throughput framework to detect synapses in electron microscopy images. *J Bioinform*. 2013;29(13):9–17.
- Jagadeesh V, Anderson J, Jones B, Marc R, Fisher S, Manjunath BS. Synapse classification and localization in Electron Micrographs. *Pattern Recognit Lett*. 2014;43(1):17–24.
- Kreshuk A, Straehle CN, Sommer C, Koethe U, Knott G, Hamprecht FA. Automated segmentation of synapses in 3D EM data. In: *The IEEE International Symposium on Biomedical Imaging*. Chicago: IEEE; 2011. p. 220–3.
- Becker C, Ali K, Knott G, Fua P. Learning context cues for synapse segmentation. *IEEE Trans Med Imaging*. 2013;32(10):1864–77.
- Kreshuk A, Koethe U, Pax E, Bock DD, Hamprecht FA. Automated Detection of Synapses in Serial Section Transmission Electron Microscopy Image Stacks. *PLoS ONE*. 2014;9(2):e87351.
- Harris KM, Perry E, Bourne J, Feinberg M, Ostroff L, Hurlburt J. Uniform serial sectioning for transmission electron microscopy. *J Neurosci*. 2006;26(47):12101–3.
- Sun M, Zhang D, Guo H, Deng H, Li W, Xie Q. 3D-reconstruction of synapses based on EM images. In: *International Conference on Materials Applications and Engineering*. Harbin: IEEE; 2016. p. 1959–1964.
- Roberts M, Jeong WK, Vazquez-Reina A, Unger M, Bischof H, Lichtman J, Pfister H. Neural process reconstruction from sparse user scribbles. In: *Medical Image Computing and Computer Assisted Interventions Conference*. Toronto: Springer; 2011. p. 621–628.
- Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. In: *Conference and Workshop on Neural Information Processing Systems*. Lake Tahoe: NIPS Foundation; 2012. p. 2852–2860.
- Beier T, Pape C, Rahaman N, et al. Multicut brings automated neurite segmentation closer to human performance. *Nat Methods*. 2017;14(2):101–2.

22. Rao Q, Xiao C, Han H, Chen X, Shen L, Xie Q. Deep learning and shapes similarity for joint segmentation and tracing single neurons in SEM images. In: SPIE Medical Imaging. Orlando: SPIE; 2017. p. 1013329.
23. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer Assisted Interventions Conference. Munich: Springer; 2015. p. 234–241.
24. Roncal WG, Pekala M, Kaynigfittkau V, et al. VESICLE: Volumetric Evaluation of Synaptic Interfaces using Computer vision at Large Scale. In: British Machine Vision Conference. Swansea: Elsevier; 2015. p. 81.1–81.13.
25. Staffler B, Berning M, Boergens KM, Gour A, Van dSP, Helmstaedter M. SynEM, automated synapse detection for connectomics. *Elife*. 2017;6: e26414.
26. Dorkenwald S, Schubert PJ, Killinger MF, et al. Automated synaptic connectivity inference for volume electron microscopy. *Nat Methods*. 2017;14(4):435–42.
27. Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transac Pattern Anal Mach Intell*. 2017;39(6):1137–49.
28. Dial R, Glover F, Karney D. Shortest path forest with topological ordering: An algorithm description in SDL. *Transp Res Part B Methodol*. 1980;14(4): 343–7.
29. Rother C, Kolmogorov V, Blake A. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Trans Graph*. 2004;23(3):309–14.
30. Schmid B, Schindelin J, Cardona A, Longair M, Heisenberg M. A high-level 3D visualization API for Java and ImageJ. *BMC Bioinformatics*. 2010;11(1):274–80.
31. Liu C, Yuen J, Torralba A, Sivic J, Freeman WT. SIFT Flow: Dense Correspondence across Different Scenes. In: European Conference on Computer Vision. Marseille: Springer; 2008. p. 28–42.
32. Schaefer S, McPhail T, Warren J. Image deformation using moving least squares. In: ACM Transactions on Graphics (TOG). Boston: ACM; 2006. p. 533–40.
33. Lia X, Jia G, Chen X, Ding W, Sun L, Xua W, Han H, Sun F. Large scale three-dimensional reconstruction of an entire *Caenorhabditis elegans* larva using AutoCUTS-SEM. *J Struct Biol*. 2017;200:87–96.
34. Girshick R. Fast R-CNN. In: IEEE International Conference on Computer Vision. Santiago: IEEE; 2015. p. 1440–1448.
35. Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. Caffe: Convolutional architecture for fast feature embedding. arXiv preprint arXiv:1408.5093. 2014.
36. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. In: European Conference on Computer Vision. Zurich: Springer; 2014. p. 818–833.
37. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In: International Conference on Learning Representations. Banff: Springer; 2014. p. 1–14.
38. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE; 2016. p. 770–8.
39. Kandel ER, Schwartz JH, Jessell TM. Principles of Neural Science, 4th Edn. McGraw-Hill xli. 2000;50(6):823–39.
40. Stauffer C, Grimson WEL. Adaptive Background Mixture Models for Real-Time Tracking. In: IEEE Conference on Computer Vision and Pattern Recognition. Fort Collins: IEEE; 1999. p. 246–52.
41. Boykov Y, Jolly M. Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images. In: IEEE International Conference on Computer Vision; 2001. p. 105–12.
42. Lucchi A, Smith K, Achanta R, Knott G, Fua P. Supervoxel- Based Segmentation of Mitochondria in EM Image Stacks with Learned Shape Features. *IEEE Transac Med Imaging*. 2012;31(2):474–86.
43. Li W, Deng H, Rao Q, Xie Q, Chen X, Han H. An automated pipeline for mitochondrial segmentation on ATUM-SEM stacks. *J Bioinform Comput Biol*. 2017;3:750015.
44. Slossberg R, Wetzler A, Kimmel R. Deep Stereo Matching with Dense CRF Priors. arXiv:1612.01725v2. 2016.
45. Çiçek Ö, Abdulkadir A, Lienkamp SS, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens: Springer; 2016. p. 424–32.
46. Dou Q, Chen H, Jin Y, et al. 3d deeply supervised network for automatic liver segmentation from ct volumes. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens: Springer; 2016. p. 149–157.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

