

Automated Detection and Segmentation of Synaptic Contacts in Nearly Isotropic Serial Electron Microscopy Images

Anna Kreshuk¹, Christoph N. Straehle¹, Christoph Sommer¹, Ullrich Koethe¹, Marco Cantoni², Graham Knott², Fred A. Hamprecht^{1*}

¹ Interdisciplinary Center for Scientific Computing, University of Heidelberg, Heidelberg, Germany, ² École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Abstract

We describe a protocol for fully automated detection and segmentation of asymmetric, presumed excitatory, synapses in serial electron microscopy images of the adult mammalian cerebral cortex, taken with the focused ion beam, scanning electron microscope (FIB/SEM). The procedure is based on interactive machine learning and only requires a few labeled synapses for training. The statistical learning is performed on geometrical features of 3D neighborhoods of each voxel and can fully exploit the high z-resolution of the data. On a quantitative validation dataset of 111 synapses in 409 images of 1948×1342 pixels with manual annotations by three independent experts the error rate of the algorithm was found to be comparable to that of the experts (0.92 recall at 0.89 precision). Our software offers a convenient interface for labeling the training data and the possibility to visualize and proofread the results in 3D. The source code, the test dataset and the ground truth annotation are freely available on the website <http://www.ilastik.org/synapse-detection>.

Citation: Kreshuk A, Straehle CN, Sommer C, Koethe U, Cantoni M, et al. (2011) Automated Detection and Segmentation of Synaptic Contacts in Nearly Isotropic Serial Electron Microscopy Images. PLoS ONE 6(10): e24899. doi:10.1371/journal.pone.0024899

Editor: Steven Barnes, Dalhousie University, Canada

Received: June 7, 2011; **Accepted:** August 19, 2011; **Published:** October 21, 2011

Copyright: © 2011 Kreshuk et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors gratefully acknowledge partial funding by the Heidelberg Graduate School for Mathematical and Computational Methods for the Sciences (<http://mathcomp.uni-heidelberg.de/>), by the Excellence Cluster Cellular Networks (<http://www.cellnetworks.uni-hd.de/>), by the ViroQuant (<http://www.viroquant.uni-hd.de/>) and SBCancer (<http://www.dkfz.de/en/sbcancer/index.html>) research initiatives, the Robert Bosch GmbH (http://www.bosch.com/worldsite_startpage/en/default.aspx) and the German Research Foundation (DFG, <http://www.dfg.de/en/index.jsp>) under grant no. HA-4364/6-1. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding was received for this study.

Competing Interests: The authors have read the journal's policy and have the following conflicts. This study was partially funded by Robert Bosch GmbH. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: fred.hamprecht@iwr.uni-heidelberg.de

Introduction

The chemical synapse is the predominant means by which information is transferred and stored in the central nervous system. Analysis of synapse size, shape and distribution contributes essential information to the understanding of neural circuitry, its function and its plasticity.

Despite the advances in light microscopy, detailed structural analysis of synapses is still only possible with electron microscopy. With serial section transmission electron microscopy (ssTEM), synaptic density can be estimated by manually counting synapses within a large volume, or by stereological extrapolation from paired 2D images [1–4]. However, using fairly thick 2D slices severely impedes detection of synapses in cases when the synaptic cleft is oriented at a low angle with respect to the plane of imaging [5].

The recent introduction of focused ion beam/scanning electron microscopy (FIB/SEM)[6] with isotropic resolution approaching 5 nm has now opened the door to a direct detection and segmentation of all synapses in large volumes of tissue, without the need to resort to extrapolation from paired slices. When searching for synapses, the human observer is not limited to the imaging plane projections of the volume, but can also explore the planes orthogonal to it. A protocol for manual synapse detection in

FIB/SEM data has recently been proposed in [7]. Still, even for the best quality EM images, manual detection of synapses remains a difficult, error-prone and time-consuming task, which calls for automated protocols to overcome the tedium of manual analysis.

To detect synapses in EM images, human experts follow a set of morphological criteria: the presence of the pre- or post-synaptic densities, a visible synaptic cleft and a nearby cluster of at least two vesicles. If an automated protocol was to be based on these criteria directly, it would require a segmentation of the entire volume to find the membrane apposition sites and a full segmentation of ultra-cellular structures to detect vesicles. Although the problem of automated segmentation of neural tissue has advanced significantly in recent years, it is not yet fully solved [8,9]. Also, automated segmentation of vesicles is nontrivial, especially at lower resolution, and has not received much attention in the literature. Rather than explicitly implementing the currently used criteria, machine learning allows to imitate the overall decisions of a human. The prediction rules are learned automatically from examples, provided in the form of annotated images (the training dataset). A meaningful measure of success is how well the automated predictions on a separate test set agree with those of the human.

Our contribution proposes an automated approach of this type and shows, through quantitative evaluation on a set of 111

synapses, that state-of-the-art machine learning methods can now achieve detection rates comparable to those of humans for asymmetric synapses in FIB/SEM data. Even though our approach does not explicitly implement the morphological criteria listed above, it finds enough evidence in the geometric features, extracted from a local neighborhood of each voxel, to mimic the decisions of the human expert.

In the field of neuroscience, recent influential work along these lines has focused on tracing and segmentation of neurons ([10–18]) or automated segmentation of ultracellular structures ([19,20]). In [9], automated synapse detection has been proposed in the course of a large-scale semi-automated volume reconstruction effort. However, this approach relies on correct partitioning of the entire volume into cells, which is still impossible by fully automated means. Finally, automated methods for synapses detection have already been proposed for fluorescence light microscopy [21,22]. Since these rely on fluorescent pre-labeling of all synapses, they are not applicable to EM images.

On the conceptual side, we rely on machine learning methods that are currently transforming all of image analysis. On the software side, we build on *ilastik* [23] and on our previous work, briefly described in [24]. *ilastik* (www.ilastik.org) is a freely available interactive learning and segmentation toolkit, which relies on a rich family of generic (nonlinear) image features and a robust nonlinear classifier [25] to estimate the probability of belonging to a synapse for each individual voxel. The training of the classifier by means of a pointing device (mouse or tablet pen) is fully interactive in the sense that a real-time display of the current predictions allows the user to iteratively provide more labels and hence improve the classifier performance. Once the classifier has been trained on a tiny subset of data, it can automatically classify all voxels in the volume as synapse or non-synapse. Then, all connected components of adjacent voxels with a sufficiently high probability of belonging to a synapse are aggregated into synapse candidates. Finally, a deterministic post-processing step rejects synapse candidates with implausible sizes. We provide a software bundle comprising a simple and intuitive graphical user interface for annotation, the machine learning algorithms and 3D visualization.

Results

The quantitative validation of the automated synapse detection procedure, as well as the evaluation of the human experts' error rate, was carried out on a test dataset of 111 asymmetric, presumed glutamatergic, synapses (see Materials and Methods section for details on data acquisition and gold standard generation).

For the evaluation of the error rate, a synapse candidate was considered to be a false positive, if its “ball” label from the human expert or its shape segmented by *ilastik* did not overlap with any ball in the gold standard dataset. If such an overlap was found, the corresponding gold standard ball was removed from the set of possible matches. Conversely, a false negative detection was counted, if a ball from the gold standard did not overlap with any of the synapse candidates; if such an overlap was found the corresponding synapse candidate was removed from the set of possible matches. Human errors were additionally reverified manually, to avoid assigning a detection error in case of a geometric disagreement between labelers, i.e. when two labelers labeled the same synapse at positions so far from each other, that their “ball” labels did not overlap.

Human experts

The expert which only had 4 hours to label and verify the synapses, missed 11 synapses and found 20 false positives. The

other two experts, unlimited in time, made 2 and 3 false negative and 7 and 8 false positive detections respectively. Most expert mistakes were made for different synapses, which is in line with the observations of [26] about attention-related errors of expert annotators of neurobiological images.

Automated detection

To quantitatively assess the algorithm performance and its stability with regard to the training data, four training sets were created from images acquired in the same experiment, but not overlapping with the test set. The four training sets were located in different parts of the image stack and contained approximately the same number of voxel labels. For each training set, 2–3 synapses were labeled, and for each of those synapses it was sufficient to only label it in one of the slices. Adding more labels did not improve the classification performance, as long as the already labeled set represented the data well, which can be judged, for example, by looking at the current algorithm predictions for some non-labeled synapses (Fig. 1, bottom row). Although the software can discriminate an arbitrary number of categories, we found three-class labeling of synapses vs. membranes vs. the rest of the tissue to produce the best results. One can also use a binary setup with synapses vs. the rest, but then the labeler has to take extra care to annotate enough membrane voxels to obtain a representative sample of the background. Adding more classes, for example, for the mitochondria, did not help the classification. Our first training set is illustrated in Fig. 1 and a performance comparison for the different training sets is shown in Fig. 2A.

After training, the classifiers were applied to the test dataset, and thresholding with different sensitivity levels was applied to the resulting synapse probability maps. Precision and recall of the algorithm, depending on the threshold, are illustrated in Fig. 2 (using the training set from Fig. 1 for Fig. 2B). Recall was calculated as the (no. of true positives)/(no. of synapses in the ground truth), precision as the (no. of true positives)/(total no. of synapse candidates). The voxelwise threshold for the detection of synaptic cores was specified as the probability of the synapse class. For the training set from Fig. 1, the best algorithm performance was at the threshold of 98%, with recall of 0.92 and precision of 0.89. Overall, the algorithm performance is better than that of a human expert working with a four-hour time limit (0.9 recall and 0.86 precision), but worse than that of domain experts with unlimited time, who, in practice, worked on the problem on two consecutive days, though not all day long (recall of 0.97 and 0.98 and precision of 0.931 and 0.936). A comparable recall value for the algorithm (0.96) was achieved at precision of 0.85. Labeling the training set, computing its appearance features and training the classifier took approximately 15 minutes. Running the algorithm on the full test dataset took several hours, however, no user interaction was needed during this time.

A 3D view of the synapses detected by the algorithm based on the training set from Fig. 1 (with probability ratio threshold of 92%) is illustrated in Fig. 3.

The human labelers only detected synapses and specified their approximate size by the ball labels, while the algorithm segmented synapses, i.e. listed every voxel belonging to a synapse candidate. Since the real synapses are not spherical, these human annotations can not serve as voxel-level gold standard. Consequently, the performance of the segmentation part of the algorithm was assessed qualitatively and found to be of sufficiently high quality for detailed analysis of synapse morphology, see Fig. 3 and Fig. 4.

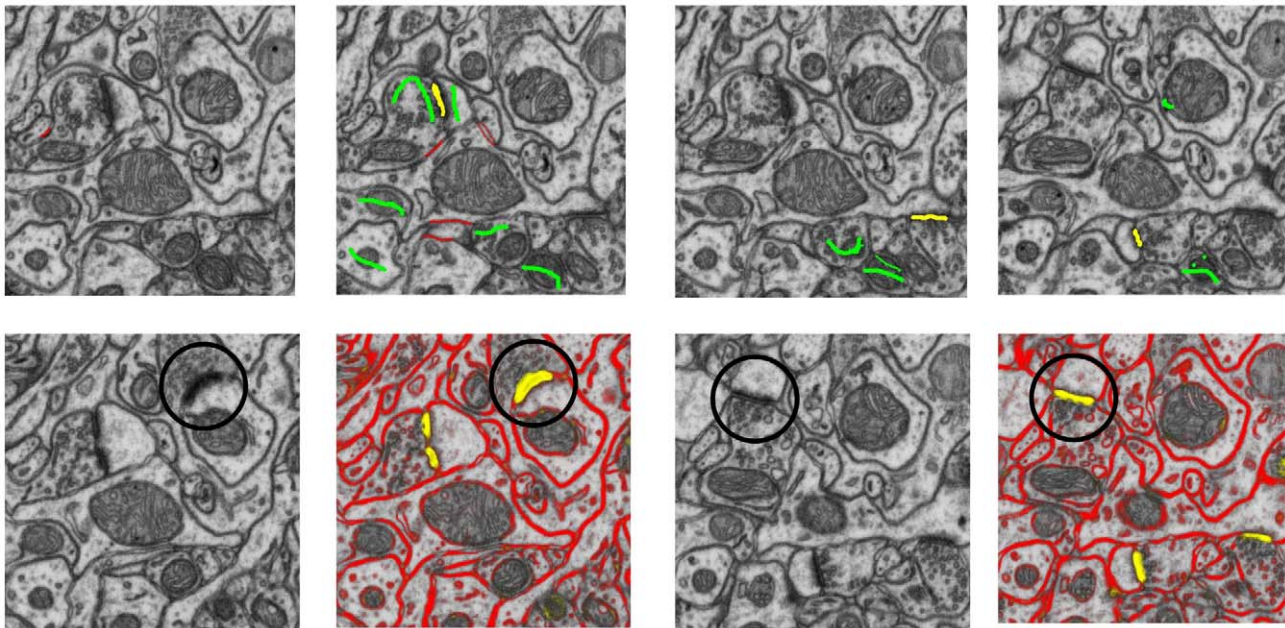


Figure 1. User labels and algorithm predictions. Top row: the complete set of user annotations for the first training set (20 brush strokes in total), with yellow labels for synapses, red for membranes, green for the rest. Bottom row: raw data and algorithm predictions on two other slices in the first training set. In black circles: some unlabeled synapses and their probability maps. The color intensity corresponds to the certainty in the prediction, predictions for green class are omitted for clarity.
doi:10.1371/journal.pone.0024899.g001

Discussion

The results show that with an adequate selection of appearance features, synapses are sufficiently different from other structures in neural tissue to allow for reliable automated detection in nearly isotropic FIB/SEM serial images. Fig. 5 illustrates typical false negative and false positive detections of the humans and of the algorithm, which have different causes. The false positives of the algorithm are mostly caused by myelinated membranes or very dark lines located near mitochondria (Fig. 5J, 5K, 5L). Similarly,

most of the false negative detections also stem from synapses located very close to myelinated membranes. In the probability maps, they become connected to the large false positives caused by these membranes, and these large connected components are then filtered out based on the size criterion (Fig. 5G). Since ilastik provides a convenient summary report of all detected synapses (Fig. 4) and reduces the data from millions of voxels to just dozens of synapse candidates, the false positives for the entire stack can easily be discarded by a human in just a few minutes of additional proofreading.

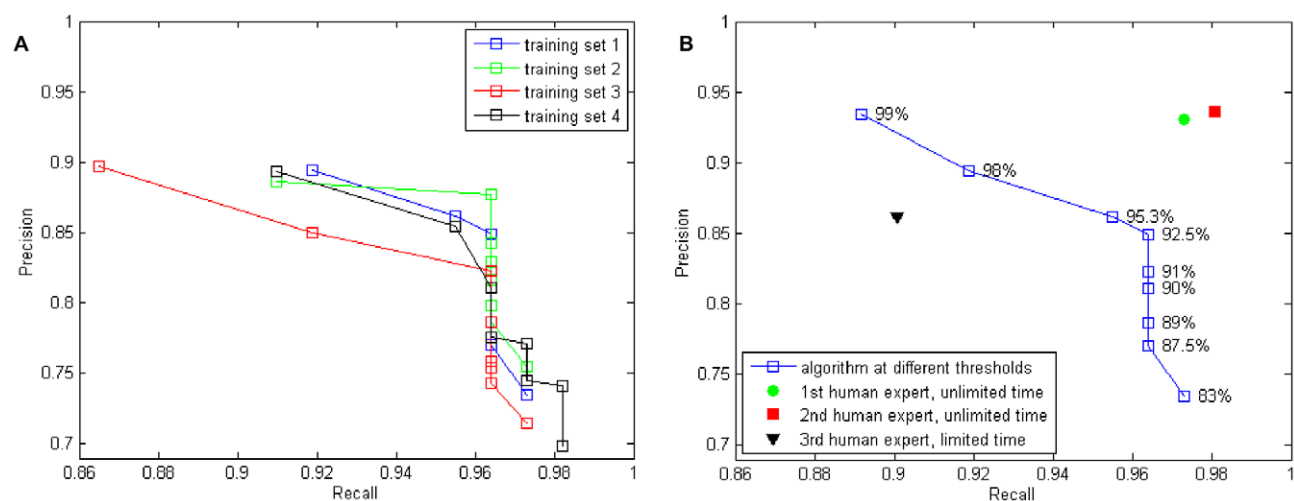


Figure 2. Precision and recall of the algorithm and the human experts. Recall was calculated as the (no. of true positives)/(no. of synapses in the ground truth), precision as the (no. of true positives)/(total no. of synapse candidates). **A:** Precision and recall of the algorithm results for the four different training sets. **B:** Precision and recall of the algorithm compared to the human experts with and without the time limit. The synapse probability threshold values are annotated next to the corresponding points of the curve.
doi:10.1371/journal.pone.0024899.g002

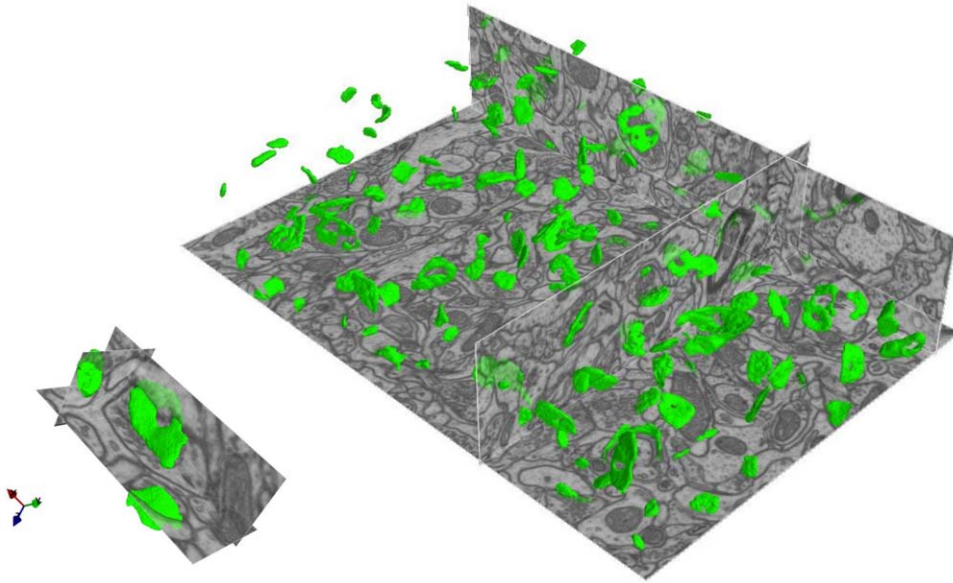


Figure 3. 3D visualization of the results. Top: all synapses detected by the algorithm after training on the labels from Fig. 1. Bottom: a close-up view of three differently oriented synapses.
doi:10.1371/journal.pone.0024899.g003

For the human experts, while some synapses that were missed are accidental omissions, others serve as a good illustration of the advantages of truly 3D processing (Fig. 5A, 5B). These synapses are oriented at a low angle to the plane of imaging and do not strictly qualify as synapses according to the morphological criteria, since the synaptic cleft is not seen in the plane of imaging. Besides that, they are just hard to discern when viewing the data in native (x-y) projection only. Since the algorithm bases its decisions on geometric features computed in full 3D neighborhoods, it is not affected by synapse orientation.

As for any machine learning-based algorithm, the performance of ilastik depends significantly on how well the training dataset represents the true variability of the test data. Note also, that the images with the training labels must be large enough to allow for computation of all features from neighborhoods of the labeled voxels. The interactive learning interface of ilastik allows the user to immediately assess the algorithm performance on a subset of data and, if necessary, to modify the training labels or the threshold value. As shown in Fig. 2A, on our data the quality of the prediction was stable with respect to the exact choice of the training set.

We expect the proposed tool to be useful not only for synapse counting, synapse density estimation or estimation of synapse-to-neuron ratio, but also for the ongoing efforts in the reconstruction of neural circuits [8,9,26–29]. We are currently working on new machine learning methods which take more spatial context into account with the aim of solving the myelinated membranes problem and achieving reliable synapse segmentation also in image stacks with low z-resolution.

Software and data availability

The software runs on Linux, MacOS and Windows. The binaries for the three platforms along with the installation instructions and documentation can be found at www.ilastik.org/synapse-detection, and the full source code is available in a github repository: [www.github.com/Ilastik/ilastik](https://github.com/Ilastik/ilastik). The test dataset, the gold standard set of synapse annotations and one of our

training label sets can also be downloaded from the website. A small downsampled test dataset is also available as part of the supporting information (Dataset S1).

Materials and Methods

Data acquisition and generation of the gold standard

The test dataset consisted of 409 scanning electron micrographs from layer 2/3 of the adult rat somatosensory cortex. The tissue preparation methods followed the protocol previously described in [6] and were performed in accordance with the procedures approved by the Office Vétérinaire Cantonale Lausanne (license number 2106). Briefly, the brain of an adult rat was fixed by cardiac perfusion of 2.5% glutaraldehyde, and 2% paraformaldehyde in phosphate buffer, it was then vibratome sectioned and slices from the somatosensory cortex were stained with buffered potassium ferrocyanide and osmium, followed by osmium, and then uranyl acetate. These stained sections were then dehydrated and embedded in Durcupan resin. The selected region was trimmed in an ultramicrotome and mounted onto an aluminium SEM stub for imaging in the FIB/SEM microscope (Zeiss NVision40), using a scanning electron beam at 1.3 kV with a current of 1 nAmp. Backscattered electrons were collected via the energy selective in-column detector (EsB) using a grid tension of 1.1 kV. The milling was achieved with a gallium ion source at 30 kV with a current of 700 pAmp. The acquired images were of 5 nm per pixel resolution with each image 1948×1342 pixels in size. The milling depth was measured at 9 nm per slice. Such high z-resolution allowed treating the data as one 3D volume of 1948×1342×409 voxels instead of a collection of 2D slices.

Synapses in the dataset were manually annotated by three independent human experts according to morphological criteria, including the presence of a pre- and post-synaptic density, as well as clustered vesicles close to the pre-synaptic membrane [30]. The human experts were researchers with experience in the analysis of electron micrographs of brain tissue and counting synapses in serial images. TrakEM2 plug-in of the FIJI framework [31] was


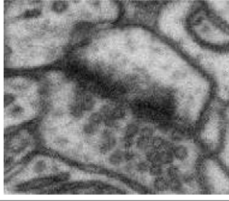
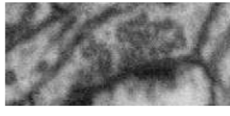
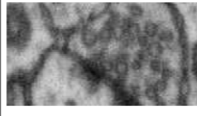

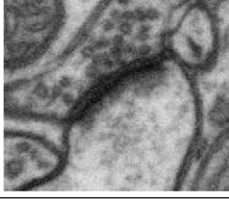

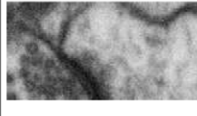


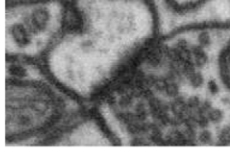
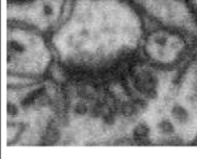


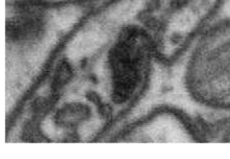
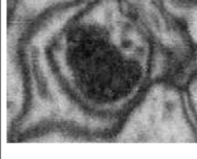

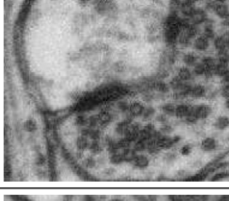
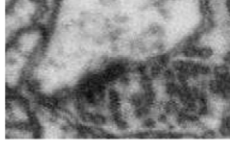
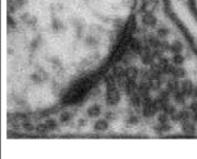





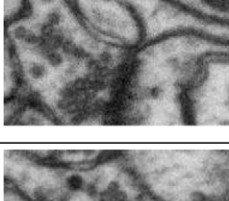
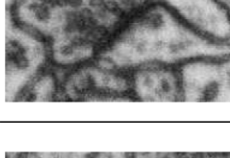
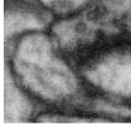

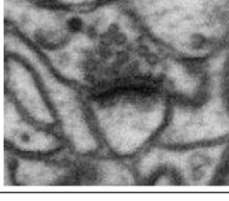

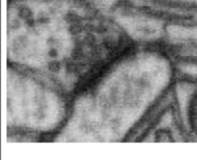
#	3d projection	XY projection	XZ projection	YZ projection	(X, Y, Z)	size in pixels
33					(666, 909, 178)	38031
34					(1087, 875, 182)	37492
35					(457, 1509, 135)	37414
36					(1070, 917, 88)	36960
37					(167, 1690, 62)	35794
38					(707, 867, 91)	34688
39					(556, 1917, 44)	33214
40					(213, 658, 59)	30086

Figure 4. Synapse detection summary report. Part of the summary report produced by ilastik. The fourth detection from the top (no. 36) is a false positive, which can easily be filtered out by a human expert by looking at a larger context.
doi:10.1371/journal.pone.0024899.g004

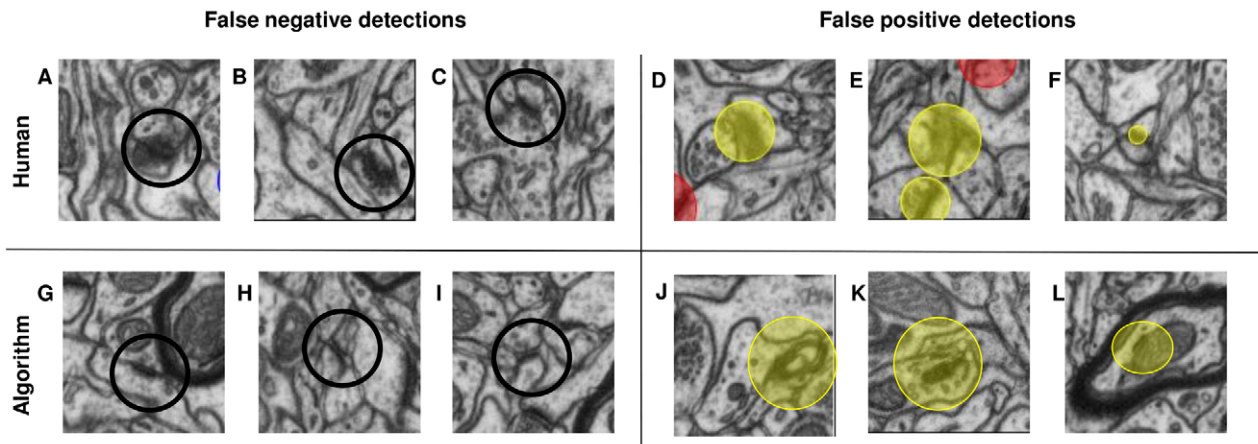


Figure 5. Error examples. A, B, C: false negative decisions of the human observers, D, E, F: false positive detections of the human observers, shown as yellow “ball” labels in the image center, G, H, I: false negative decisions of the algorithm, J, K, L false positive decisions of the algorithm. doi:10.1371/journal.pone.0024899.g005

used for the annotation. One of the experts only had four hours to label and verify the complete dataset, while the other two experts were not limited in time and took several hours longer. The annotation of each expert included positions and approximate size of detected synapses, denoted by “ball” labels from TrakEM2. Some examples of expert labels can be seen in Fig. 5D, 5E, 5F. Each expert first analyzed the dataset independently from the others and the resulting three sets of annotations were compared automatically to find all discrepancies. Since the automatic comparison procedure found differences between the expert annotations, these cases had to be re-examined jointly by all experts to establish a gold standard annotation. Synapses touching the left or top border of the image, as well as those touching the last slice of the stack, were excluded from the final count. For evaluation purposes, we also excluded synapses which had their center in the first slice of the stack, to avoid the border effects described in the next section. The resulting set of 111 synapses formed the gold standard and was used to estimate the error rates of both the original human annotations and the results obtained by the algorithm.

Algorithm

The input data for the algorithm consists of scanning electron micrographs of neural tissue, provided as a pre-registered image stack, and user labels on a tiny subset of the data. The labeling can be very sparse, as shown in Fig. 1. The standard EM protocol used to prepare the brain tissue for imaging gives high contrast not only to synapses, but also to other cellular structures, such as mitochondria. As a consequence, the classification cannot simply be based on the raw intensity values of individual voxels. Instead, more informative features are required that also encode geometrical properties of 3D voxel neighborhoods. Different features represent different properties of these neighborhoods and should be selected so as to allow for an effective discrimination of the labeled classes. For example, as synapses are darker than intracellular space, the average intensity would serve as a good feature to distinguish these two, but would not help to separate synapses from membranes or mitochondria. Edge detectors respond strongly to both membranes and endoplasmic reticulum. Texture features respond to synapses, but also pick up thick mitochondrial membranes. Rather than devise decision rules by hand, we use statistical learning from a labeled training set to infer robust classification rules.

Since features have to be computed for every voxel, memory consumption has to be taken into account for large volumes. To allow running of the algorithm on a modern desktop PC rather than a high-end server without compromising classification accuracy, we performed selection of features, based on their Gini importance [25]. The final list of 38 features is provided in Table 1. Although the user is free to re-adjust the list and try out new feature combinations, we do not expect it to be necessary, except for the adjustment of the neighborhood sizes to the resolution of the data. Due to boundary effects in the feature computation, the performance of the algorithm can decrease for voxels very close to the limits of the dataset, such as the voxels of the first and last scan of the stack.

Based on the features and user labels, the Random Forest classifier [25] computes a probability map for each voxel, i.e. its probability of belonging to one of the classes defined in the training phase. Random Forest is a bagged ensemble of randomized decision trees that has only two parameters: the number of trees and the number of features considered at each split. Random Forest has been empirically shown to be fairly robust to their choice, and to provide very good results for a broad

Table 1. Voxel features.

Feature	Sigmas	# of channels
Eigenvalues of the Hessian matrix	1, 1.6, 3.5, 5	3
Eigenvalues of the structure tensor	1, 1.6, 3.5, 5	3
Intensity of the Gaussian-smoothed image	0.7, 1, 1.6, 3.5, 5	1
Gradient magnitude of the Gaussian-smoothed image	1.6, 3.5, 5	1
Laplacian of the Gaussian-smoothed image	1.6, 3.5, 5	1
Difference of Gaussians	1.6, 3.5, 5	1
Total number of channels		38

Local neighborhood features, used for voxel classification. The “Sigmas” column shows the standard deviation of the Gaussians, used for smoothing the data. This parameter effectively determines the size of the necessary voxel neighborhood. For the eigenvalues of the structure tensor, the second scale parameter was set to $\sigma/2$, for the difference of Gaussians the second Gaussian sigma was set to $0.66 \cdot \sigma$. doi:10.1371/journal.pone.0024899.t001

range of applications [32–35]. An example of Random Forest probability maps is shown in Fig. 1 (note the soft borders of the classes, which show that it's a probability estimate, not a hard segmentation).

The obtained probability maps are smoothed by convolution with a Gaussian with a standard deviation of 5 voxels to avoid local discontinuities caused by noisy voxel-wise predictions. Uncertain detections are then filtered out by considering only those clusters of voxels with synapse probability greater than a given threshold and with size of at least 1000 voxels. The lower limit for the size filter was computed as the approximate volume occupied by two vesicles at the given data resolution. The probability threshold can be interactively adjusted by the user. After thresholding, only the cores of synapses, i.e. areas of very high synapse probability, are left. These cores underestimate the real size of synapses, so to transition from detection to a proper segmentation we relax the synapse probability threshold to 0.5 for all voxels that are adjacent to synaptic cores.

Software

The freely available *ilastik* toolkit [23] provides an intuitive interface for classification and segmentation of 2D and 3D data. In the interactive mode, it allows the user to immediately see the effect of newly added labels on the classifier's predictions, and therefore reduces the necessary labeling time. Once the classifier has been trained on a representative subset of the data, predictions on a very large dataset can be performed off-line in batch-processing mode.

References

1. Sterio DC (1984) The unbiased estimation of number and sizes of arbitrary particles using the disector. *Journal of Microscopy* 134: 127–136.
2. Geinisman Y, Gundersen HJG, Zee E, West MJ (1996) Unbiased stereological estimation of the total number of synapses in a brain region. *Journal of Neurocytology* 25: 805–819.
3. Mayhew TM (1996) How to count synapses unbiasedly and efficiently at the ultrastructural level: proposal for a standard sampling and counting protocol. *Journal of Neurocytology* 25: 793–804.
4. Coggeshall RE, Lekan HA (1996) Methods for determining numbers of cells and synapses: A case for more uniform standards of review. *The Journal of Comparative Neurology* 364: 6–15.
5. Kubota Y (2009) Important factors for the three-dimensional reconstruction of neuronal structures from serial ultrathin sections. *Frontiers in Neural Circuits* 3.
6. Knott G, Marchman H, Wall D, Lich B (2008) Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *Journal of Neuroscience* 28: 2959–2964.
7. Merchan-Perez A, Rodriguez J, Alonso-Nanclares L, Schertel A, Defelipe J (2009) Counting synapses using FIB/SEM microscopy: A true revolution for ultrastructural volume reconstruction. *Frontiers in Neuroanatomy* 3.
8. Chklovskii DB, Vitaladevuni S, Scheffer LK (2010) Semi-automated reconstruction of neural circuits using electron microscopy. *Current Opinion in Neurobiology* 20: 667–675.
9. Mishchenko Y, Hu T, Spacek J, Mendenhall J, Harris KM, et al. (2010) Ultrastructural analysis of hippocampal neuropil from the connectomics perspective. *Neuron* 67: 1009–1020.
10. Andres B, Koethe U, Helmstaedter M, Denk W, Hamprecht FA (2008) Segmentation of SBFSEM volume data of neural tissue by hierarchical classification. In: Rigoll G, editor, *Pattern Recognition*. Springer, volume 5096 of *LNC3*. pp 142–152.
11. Mishchenko Y (2009) Automation of 3D reconstruction of neural tissue from large volume of conventional serial section transmission electron micrographs. *Journal of Neuroscience Methods* 176: 276–289.
12. Vazquez-Reina A, Miller E, Pfister H (2009) Multiphase geometric couplings for the segmentation of neural processes. In: *Proceedings of CVPR*. Miami, FL. pp 2020–2027.
13. Anderson JR, Jones BW, Yang J, Shaw MV, Watt CB, et al. (2009) A computational framework for ultrastructural mapping of neural circuitry. *PLoS Biology* 7: e1000074.
14. Turaga SC, Briggman KL, Helmstaedter M, Denk W, Seung SH (2009) Maximum affinity learning of image segmentation. In: *Proceedings of NIPS*. volume abs/0911.5372.
15. Jurrus E, Tasdizen T, Koshevoy P, Fletcher PT, Hardy M, et al. (2009) Axon tracking in serial Block-Face scanning electron microscopy. *Medical Image Analysis* 13: 180–188.
16. Kaynig V, Fuchs T, Buhmann JM (2010) Neuron geometry extraction by perceptual grouping in serial TEM images. In: *Proceedings of CVPR*. volume 0. pp 2902–2909.
17. Jain V, Bollmann B, Richardson M, Berger D, Helmstaedter M, et al. (2010) Boundary learning by optimization with topological constraints. In: *Proceedings of CVPR*. pp 2488–2495.
18. Jurrus E, Paiva ARC, Watanabe S, Anderson JR, Jones BW, et al. (2010) Detection of neuron membranes in electron microscopy images using a serial neural network architecture. *Medical Image Analysis* 14: 770–783.
19. Narasimha R, Ouyang H, Gray A, McLaughlin SW, Subramaniam S (2009) Automatic joint classification and segmentation of whole cell 3D images. *Pattern Recognition* 42: 1067–1079.
20. Lucchi A, Smith K, Achanta R, Lepetit V, Fua P (2010) A fully automated approach to segmentation of irregularly shaped cellular structures in EM images. In: *Proceedings of MICCAI*. Beijing, China, .
21. Herold J, Schubert W, Nattkemper TW (2010) Automated detection and quantification of uorescently labeled synapses in murine brain tissue sections for high throughput applications. *Journal of Biotechnology* 149: 299–309.
22. Schmitz SK, Hjorth JJJ, Joemai RM, Wijntjes R, Eijgenraam S, et al. Automated analysis of neuronal morphology~ synapse number and synaptic recruitment. *Journal of Neuroscience Methods* In Press, Accepted Manuscript. ? ? ?-1.
23. Sommer C, Strachle CN, Koethe U, Hamprecht FA (2011) *ilastik*: interactive learning and segmentation toolkit. In: *Proceedings of ISBI*. Chicago, IL, USA.
24. Kreshuk A, Strachle CN, Sommer C, Koethe U, Knott G, et al. (2011) Automated segmentation of synapses in 3D EM data. In: *Proceedings of ISBI*. Chicago, IL, USA.
25. Breiman L (2001) Random forests. *Machine Learning* 45: 5–32.
26. Helmstaedter M, Briggman KL, Denk W (2011) High-accuracy neurite reconstruction for high-throughput neuroanatomy. *Nature Neuroscience* 14: 1081–1088.
27. Briggman KL, Denk W (2006) Towards neural circuit reconstruction with volume electron microscopy techniques. *Current Opinion in Neurobiology* 16: 562–570.
28. Helmstaedter M, Briggman KL, Denk W (2008) 3D structural imaging of the brain with photons and electrons. *Current Opinion in Neurobiology* 18: 633–641.
29. Jain V, Seung HS, Turaga SC (2010) Machines that learn to segment images: a crucial technology for connectomics. *Current Opinion in Neurobiology* 20: 653–666.
30. Knott GW, Quairiaux C, Genoud C, Welker E (2002) Formation of dendritic spines with GABAergic synapses induced by whisker stimulation in adult mice. *Neuron* 34: 265–273.

31. Cardona A, Saalfeld S, Preibisch S, Schmid B, Cheng A, et al. (2010) An Integrated Micro- and Macroarchitectural Analysis of the *Drosophila* Brain by Computer-Assisted Serial Section Electron Microscopy. *PLoS Biol* 8: e1000502.
32. Guo L, Ma Y, Cukic B, Singh H (2004) Robust prediction of fault-proneness by random forests. In: *Proceedings of ISSRE*. pp 417–428.
33. Caruana R, Niculescu-Mizil A (2006) An empirical comparison of supervised learning algorithms. In: *Proceedings of ICML*. Pittsburgh, Pennsylvania: ACM. pp 161–168.
34. Diaz-Uriarte R, de Andres SA (2006) Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* 7: 3.
35. Menze BH, Kelm BM, Masuch R, Himmelreich U, Bachert P, et al. (2009) A comparison of random forest and its gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics* 10: 213–213.
36. Schroeder W, Martin K, Lorensen B (2006) *The Visualization Toolkit*. <http://www.vtk.org>, fourth edition.