



Synaptic Cleft Segmentation in Non-isotropic Volume Electron Microscopy of the Complete *Drosophila* Brain

Larissa Heinrich¹, Jan Funke^{1,2}, Constantin Pape^{1,3}, Juan Nunez-Iglesias¹,
and Stephan Saalfeld¹(✉)

¹ HHMI Janelia Research Campus, Ashburn, USA

saalfelds@janelia.hhmi.org

² Institut de Robòtica i Informàtica Industrial, Barcelona, Spain

³ University of Heidelberg, Heidelberg, Germany

Abstract. Neural circuit reconstruction at single synapse resolution is increasingly recognized as crucially important to decipher the function of biological nervous systems. Volume electron microscopy in serial transmission or scanning mode has been demonstrated to provide the necessary resolution to segment or trace all neurites and to annotate all synaptic connections.

Automatic annotation of synaptic connections has been done successfully in near isotropic electron microscopy of vertebrate model organisms. Results on non-isotropic data in insect models, however, are not yet on par with human annotation.

We designed a new 3D-U-Net architecture to optimally represent isotropic fields of view in non-isotropic data. We used regression on a signed distance transform of manually annotated synaptic clefts of the CREMI challenge dataset to train this model and observed significant improvement over the state of the art.

We developed open source software for optimized parallel prediction on very large volumetric datasets and applied our model to predict synaptic clefts in a 50 tera-voxels dataset of the complete *Drosophila* brain. Our model generalizes well to areas far away from where training data was available.

1 Introduction

Today, the neuroscience community widely agrees that the synaptic microcircuitry of biological nervous systems is important to understand what functions they implement. The only currently available method to densely reconstruct all axons, dendrites, and synapses is volume electron microscopy (EM) as it provides a resolution sufficient to unambiguously separate them (<15 nm per voxel, [19]). For EM *connectomics*, several flavors of volume EM have been used [3]: Serial block face scanning EM (SBFSEM), in combination with focused ion beam milling (FIB-SEM), provides the highest isotropic resolution of $\sim 5^3$ nm per voxel

and excellent signal to noise ratio but is relatively slow. On the other end of the spectrum, serial section transmission EM (ssTEM) offers excellent lateral resolution, imaging speed, and signal to noise ratio but generates highly non-isotropic data with comparably poor axial resolution (>35 nm per voxel).

A remarkable number of projects are currently under way to reconstruct the connectomes of various model organisms [23], ranging from small invertebrate nervous systems like the larvae of *Drosophila melanogaster* [6, 24] or *Platynereis dumerilii* [20], the adult *Drosophila* [27, 28], to vertebrate models like the zebrafish larva [9], the retina of a mouse [8], or the zebra finch HVC [13].

2 Related Work

While many ongoing connectome reconstruction efforts still rely on manual annotation of synaptic contacts [24, 28], automatic annotation of synaptic clefts from volume electron microscopy has been explored in recent years. On vertebrate model systems, existing solutions perform comparably to trained human annotators on both isotropic [2, 5, 14, 15, 26] and non-isotropic data [5, 16]. Synapses in the insect brain, however, are more complicated, and typically smaller than in vertebrates. Accordingly, the performance on isotropic data is good [10, 14, 18], but not yet satisfying on non-isotropic data (see CREMI leaderboard).¹

The methods follow the general trend in computer vision. Earlier approaches [2, 14–16, 18] use carefully designed image features and train pixel classifiers using random forests or gradient boosting. More recent approaches [5, 10, 26] train deep learning models to classify pixels or regions of interest as synapse candidates. All approaches rely on sensible post-processing to filter false detections.

The CREMI challenge provides three volumes with ground truth for neuron segmentation, synaptic clefts, and synaptic partner annotations in diverse regions of ssTEM of the adult *Drosophila* brain at 40×4^2 nm per voxel. The challenge data includes typical artifacts for ssTEM preparations such as missing sections, staining precipitate, or incorrect alignment. To our knowledge, it is the only existing challenge with secret test data that enables unbiased comparison of synapse detection in non-isotropic EM of the insect brain. The evaluation metric for synaptic cleft detection (CREMI score) is the average of the average false positive distance (FPD) and the average false negative distance (FND). The FPD is the distance of a predicted label to the nearest true label, the FND is the distance of a true label to the nearest predicted label.

3 Methods

3.1 Training Setup

We corrected the serial section alignment errors present in the CREMI volumes using elastic alignment with TrakEM2 [22] and split each volume into a training (75%) and validation (25%) subset, such that the statistics of each subset

¹ MICCAI Challenge on Circuit Reconstruction from Electron Microscopy Images (CREMI): <https://cremi.org>.

are visually similar to the whole block. We trained 3D-U-Nets [4] to predict a signed distance transform of binary synapse labels using TensorFlow [1]. We used Gunpowder² for batch loading, preprocessing, and training. We made heavy use of Gunpowder’s support for data augmentations such as transposing, intensity variation, elastic deformations, and ssTEM-specific artifacts like missing or noisy sections. We believe that these augmentations are crucial for our network to generalize well on large datasets without substantial engineering efforts.

As synaptic clefts are very sparse, we sample batches that contain synapses more frequently by rejecting batches without synapses with 95% probability. Additionally, we rebalance the loss with the frequency of positively annotated voxels to heavily penalize false negative predictions (unless otherwise stated).

We used Adam to minimize the L2 loss w.r.t. a signed Euclidean distance transform (SEDT) of the binary labels. As the SEDT is not meaningful far away from synapses, we scaled it and applied a tanh nonlinearity that saturates between $[-1, 1]$: $STDT = \tanh(SEDT/s)$. Our experiments indicated that the scaling factor has little effect on performance (data not shown), we chose $s = 50$. Simple thresholding converts the predicted $STDT$ into binary labels.

3.2 Experiments

3D-U-Nets Benefit from Isotropic Fields of View (FOV). The DTU-1 (distance transform U-Net) architecture is based on a design for neuron segmentation in non-isotropic ssTEM [7] (Fig. 1). The FOV of this architecture is highly non-isotropic across a large number of layers. Hypothesizing that an isotropic FOV would be beneficial to learn meaningful physical features, we tweaked the kernel sizes while retaining the overall design. The ‘isotropic’ network (DTU-2, Fig. 1) is restricted to 2D convolutions in the first few levels and has isotropic kernels once the voxel size is nearly isotropic. The encoding and decoding side are symmetric. Figure 2 shows that DTU-2 consistently outperforms DTU-1.

DTU-2 has significantly fewer parameters than DTU-1. This allows for a larger patch size (output size $23 \times 218 \times 218$ as opposed to $56 \times 56 \times 56$) which translates into a better estimate of the true gradient during stochastic gradient descent. While this constitutes an additional advantage of DTU-2, we showed that it is not sufficient to explain its superior performance. A smaller version of DTU-2, with output size $20 \times 191 \times 191$, still outperforms DTU-1 (Fig. 2a).

At the time of writing, DTU-2 is first on the CREMI synaptic cleft detection challenge, followed by DTU-1 in second place. Unlike the experiments shown in Fig. 2, those networks were trained for more iterations, on a curated version of the full CREMI ground truth.

Regression Outperforms Classification. Most deep-learning approaches for object detection in general, and synapse detection specifically [5], use a sigmoid nonlinearity and cross-entropy loss to predict a probability map for the object. Inspired by the recent success of long-range affinities as an auxiliary loss for

² Gunpowder: <https://github.com/funkey/gunpowder>.

boundary detection [17], we suspected that networks might generally benefit from being explicitly forced to gather information from a larger context. With this assumption in mind, we trained the network to predict a distance rather than a probability map. This approach turns the voxel-wise classification into a voxel-wise regression problem [12].

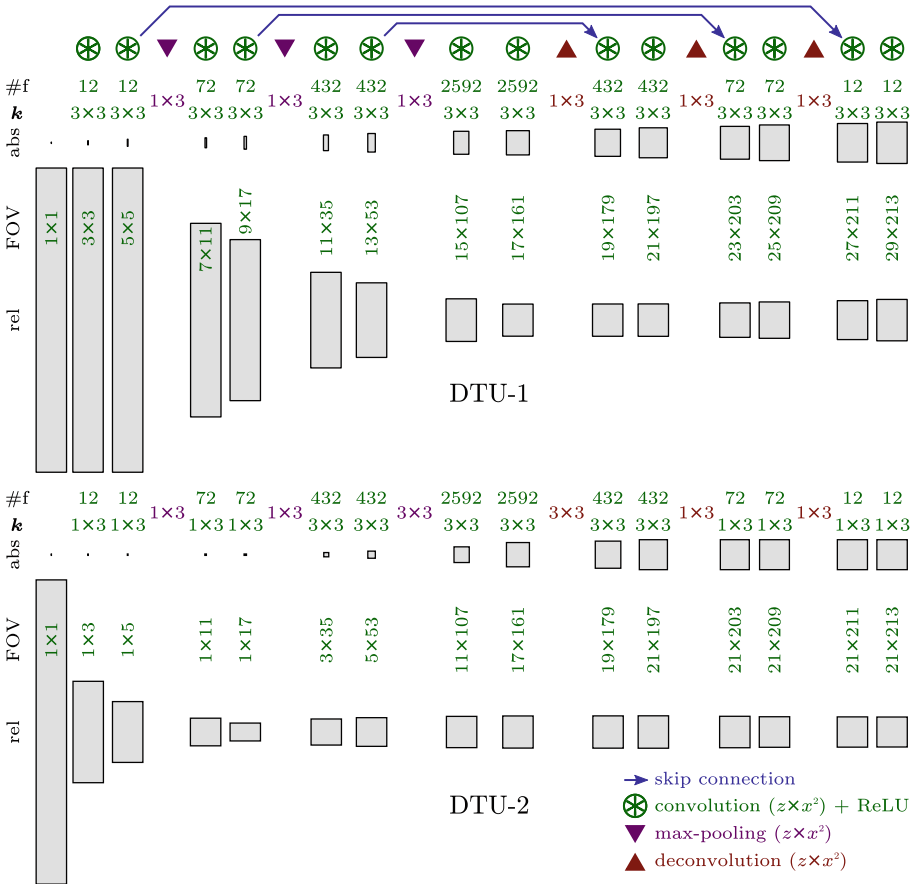


Fig. 1. Comparison of the physical FOV in each layer of the 3D-U-Net architectures DTU-1 and DTU-2. The top row shows a graphical representation of the general U-Net architecture. The network consists of modules of two convolutions and rectilinear units followed by max-pooling on the encoding side, and deconvolution followed by two convolutions and rectilinear units on the decoding side. Kernel sizes (k) are denoted as $z \times x$ as the x and y axes are isotropic. The number of features per convolutional layer ($\#f$) is increased by a factor of six after max-pooling and decreased by a factor of six after deconvolution. In DTU-2, 3D-convolutions are replaced by 2D-convolutions where the resolution is highly non-isotropic, and 2D-max-pooling and deconvolution are replaced by 3D-max-pooling and deconvolution where the resolution is near-isotropic. The physical FOV in each layer, depicted as absolute (abs) and relative (rel) size boxes, is therefore closer to isotropic than in DTU-1.

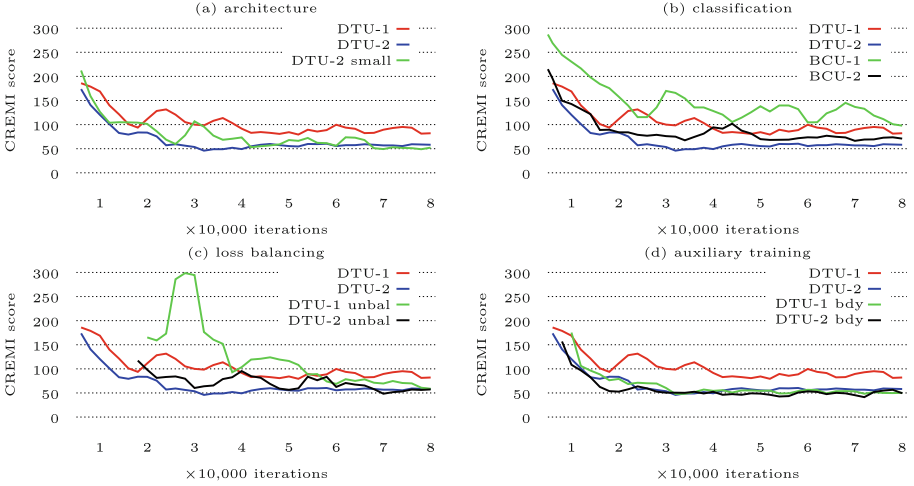


Fig. 2. Validation experiments. All plots show mildly smoothed validation results sampled in intervals of 2,000 iterations using the CREMI score averaged over the validation set. (a) shows that DTU-2 outperforms DTU-1, even if training blocks of the same size as for DTU-1 are used. (b) shows that DTU-1 and DTU-2 trained for regression on the distance transform outperform the same architectures trained for binary classification. (c) shows that loss balancing makes training more robust. (d) shows that auxiliary training for boundary distances improves performance on synaptic cleft detection.

In Fig. 2b, we compare the performance of probability map prediction using a sigmoid nonlinearity with binary cross entropy loss and the STDT prediction as shown before. All other hyperparameters are the same and the maps are converted into binary labels with a non-tweaked threshold (i.e. 0.5 and 0, respectively). For both network architectures, the CREMI score on the validation set improves when predicting the STDT.

Loss Balancing Is Important. Rebalancing the loss is an important feature of the training pipeline (Fig. 2c). In early iterations, the CREMI score cannot be properly evaluated as no voxel in the validation set is predicted to be above threshold, i.e. no synapses were detected.

Auxiliary Training Improves Performance. As synaptic clefts are, by definition, located at cell boundaries, we conducted experiments to determine whether an auxiliary loss from predicting a distance map of cell boundaries boosts performance. We added a second output channel to both DTU-1 and DTU-2 with an (unbalanced) L2 loss with respect to the STDT, now computed on the neuron labels. The two losses are weighed equally. Batch sampling is still done with respect to synaptic clefts.

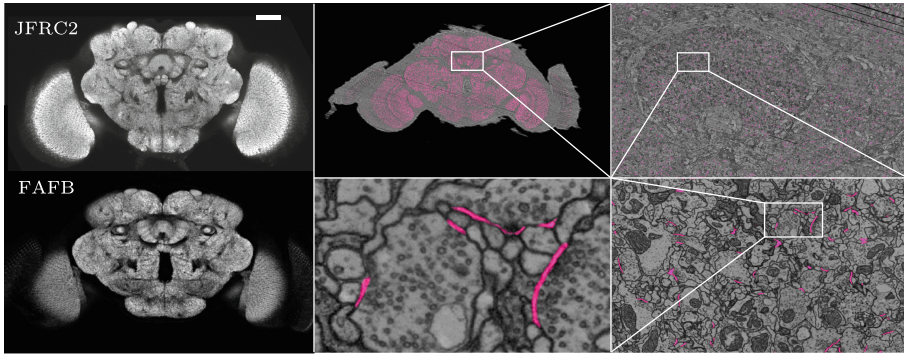


Fig. 3. Synaptic cleft prediction on the complete adult *Drosophila* brain. Left: Convolution of our predictions (FAFB) with a smooth PSF reproduces synaptic densities as visualized by fluorescent microscopy with the the nc82 antibody (JFRC2 template brain [11]), scale bar 50 μm . Right: Exemplary zoom-series into our DTU-2 synaptic cleft predictions overlaid over the FAFB volume.

Figure 2d shows that both networks benefit from the auxiliary loss signal. Interestingly, the effect is more significant for DTU-1. A careful evaluation of the boundary detection is beyond the scope of this work.

3.3 Synaptic Cleft Prediction on the Complete *Drosophila* Brain

Prediction on large volumes can be performed in parallel on adjacent blocks. Since our network was trained on valid input, the input block-size needs to be padded by the FOV of the network, i.e. while output blocks are adjacent and non-overlapping, input blocks overlap.

We converted the full adult fly brain (FAFB) volume [28]³ into a scale-pyramid using the the N5 data format⁴ on a shared filesystem. We used N5 for both input and output because it enables parallel reading and writing of compressed volumetric blocks. Prediction requires less memory than training because gradients do not need to be computed. We found $71 \times 650 \times 650$ voxels to be the maximum valid output block-size for the DTU-2 network that we could process on our NVIDIA Quadro M6000 GPUs with 12 GB of RAM. Using this increased block-size accelerated prediction by a factor of ~ 2.5 compared to the block-size used for training.

The relevant biological sample covers only 20% of the FAFB volume. We used ilastik [25] to train a random forest classifier on scale-level 7 (downscaled by $13 \times 128 \times 128$, i.e. $\sim 0.5^3 \mu\text{m}$ per voxel) that separates relevant biological sample from background. Only output blocks that intersect with this mask were considered for prediction. This valid set of blocks has a volume of ~ 50 tera-voxels, the entire FAFB volume including background contains ~ 213 tera-voxels.

³ Available for download at <http://temca2data.org/>.

⁴ N5 specification: <https://github.com/saalfeldlab/n5>.

We distributed the list of output blocks over 48 GPUs. For each GPU, we used Dask [21] to load, preprocess, and store image blocks while the GPU performed prediction, achieving greater than 90% GPU utilization⁵. Our average prediction speed was ~ 3 mega-voxels per second and GPU, i.e. prediction of the complete volume was finished in less than five days.

The quality of predictions across the entire volume was consistent with our results on CREMI (see Fig. 3). Even in areas with different characteristics than the CREMI training volumes (such as the lamina), synaptic cleft predictions are mostly correct and consistent with our expectations. Predictions are correctly missing in axonal tracts and in the cortex. We produced a simulation of an nc82 labeled confocal image by applying a large non-isotropic Gaussian PSF to our predictions and visually compared the result with the JFRC2 template brain [11] (Fig. 3). Accounting for that the two volumes stem from different individuals and have not been registered, our predictions convincingly reproduce the synaptic density distribution as visualized with the nc82 antibody.

4 Conclusion

We achieved a significant improvement over the state of the art in detection and segmentation of synaptic clefts in non-isotropic ssTEM of the insect nervous system. We designed a 3D-U-Net architecture and training scheme that is particularly well suited to account for the non-isotropy in ssTEM data and the sparsity of synapses. We trained this architecture by regression on a signed distance transform of manually annotated synaptic clefts of the publicly available CREMI challenge. We showed that our new architecture compares favorably to a previously described architecture for the same data despite exposing fewer training parameters. We developed an optimized framework for parallel prediction on very large volumetric data and achieved a prediction throughput of ~ 3 mega-voxels per second and GPU. This efficiency enabled us to predict all synaptic clefts in the 50 tera-voxels full adult *Drosophila* brain [28] in less than five days. We made our code available as open source under a permissive license⁶.

References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., et al.: TensorFlow: large-scale machine learning on heterogeneous systems (2015)
2. Becker, C., Ali, K., Knott, G., Fua, P.: Learning context cues for synapse segmentation in EM volumes. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012. LNCS, vol. 7510, pp. 585–592. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33415-3_72
3. Briggman, K.L., Bock, D.D.: Volume electron microscopy for neuronal circuit reconstruction. *Curr. Opin. Neurobiol.* **22**(1), 154–161 (2012)

⁵ Parallel prediction framework: <https://github.com/saalfeldlab/simpleference>.

⁶ CNNectome: <https://github.com/saalfeldlab/cnnectome>.

4. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
5. Dorkenwald, S., Schubert, P.J., Killinger, M.F., Urban, G.: Automated synaptic connectivity inference for volume electron microscopy. *Nat. Methods* **14**, 435–442 (2017)
6. Eichler, K., Li, F., Litwin-Kumar, A., Park, Y., et al.: The complete connectome of a learning and memory centre in an insect brain. *Nature* **548**, 175–182 (2017)
7. Funke, J., Tschopp, F., Grisaitis, W., Sheridan, A., et al.: Large scale image segmentation with structured loss based deep learning for connectome reconstruction. *IEEE Trans. Pattern Anal. Mach. Intell.* (early access). <https://doi.org/10.1109/TPAMI.2018.2835450>
8. Helmstaedter, M., Briggman, K.L., Turaga, S.C., Jain, V., et al.: Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature* **500**, 168–174 (2013)
9. Hildebrand, D.G.C., Cicconet, M., Torres, R.M., Choi, W., et al.: Whole-brain serial-section electron microscopy in larval zebrafish. *Nature* **545**, 345–349 (2017)
10. Huang, G.B., Plaza, S.: Identifying synapses using deep and wide multiscale recursive networks. [arXiv:1409.1789](https://arxiv.org/abs/1409.1789) [cs.CV] (2014)
11. Jenett, A., Rubin, G., Ngo, T.T., Shepherd, D., et al.: A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* **2**(4), 991–1001 (2012)
12. Kainz, P., Urschler, M., Schultze, S., Wohlhart, P., Lepetit, V.: You should use regression to detect cells. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 276–283. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_33
13. Kornfeld, J., Benezra, S.E., Narayanan, R.T., Svava, F., et al.: EM connectomics reveals axonal target variation in a sequence-generating network. *eLife* **6**, e24364 (2017)
14. Kreshuk, A., Straehle, C., Sommer, C., Koethe, U., et al.: Automated detection and segmentation of synaptic contacts in nearly isotropic serial electron microscopy images. *PloS One* **6**(10), e24899 (2011)
15. Kreshuk, A., Funke, J., Cardona, A., Hamprecht, F.A.: Who is talking to whom: synaptic partner detection in anisotropic volumes of insect brain. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9349, pp. 661–668. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24553-9_81
16. Kreshuk, A., Koethe, U., Pax, E., Bock, D.D., Hamprecht, F.A.: Automated detection of synapses in serial section transmission electron microscopy image stacks. *PLOS ONE* **9**(2), 1–11 (2014)
17. Lee, K., Zung, J., Li, P., Jain, V., Seung, H.S.: Superhuman accuracy on the SNEMI3D connectomics challenge. [arXiv:1706.00120](https://arxiv.org/abs/1706.00120) [cs.CV] (2017)
18. Plaza, S., Parag, T., Huang, G., Olbris, D., et al.: Annotating synapses in large EM datasets. [arXiv:1409.1801v2](https://arxiv.org/abs/1409.1801v2) (2014)
19. Plaza, S.M., Scheffer, L.K., Chklovskii, D.B.: Toward large-scale connectome reconstructions. *Curr. Opin. Neurobiol.* **25**, 201–210 (2014)
20. Randel, N., Shahidi, R., Verasztó, C., Bezares-Calderón, L.A., et al.: Inter-individual stereotypy of the *Platynereis* larval visual connectome. *eLife* **4**, e08069 (2015)
21. Rocklin, M.: Dask: parallel computation with blocked algorithms and task scheduling. In: *Python in Science*, pp. 130–136 (2015)

22. Saalfeld, S., Fetter, R., Cardona, A., Tomancak, P.: Elastic volume reconstruction from series of ultra-thin microscopy sections. *Nat. Methods* **9**(7), 717–720 (2012)
23. Schlegel, P., Costa, M., Jefferis, G.S.: Learning from connectomics on the fly. *Curr. Opin. Insect Sci.* **24**, 96–105 (2017)
24. Schneider-Mizell, C.M., Gerhard, S., Longair, M., Kazimiers, T., et al.: Quantitative neuroanatomy for connectomics in *Drosophila*. *eLife* **5**, e12059 (2016)
25. Sommer, C., Straehle, C., Köthe, U., Hamprecht, F.A.: Ilastik: interactive learning and segmentation toolkit. In: ISBI, pp. 230–233 (2011)
26. Staffler, B., Berning, M., Boergens, K.M., Gour, A.: SynEM, automated synapse detection for connectomics. *eLife* **6**, e26414 (2017)
27. Takemura, S.Y., Aso, Y., Hige, T., Wong, A.: A connectome of a learning and memory center in the adult *Drosophila* brain. *eLife* **6**, e26975 (2017)
28. Zheng, Z., Lauritzen, J.S., Perlman, E., Robinson, C.G., et al.: A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*. *Cell* **174**(3), 730–743.e22 (2018). <https://doi.org/10.1016/j.cell.2018.06.019>