

NucMM Dataset: 3D Neuronal Nuclei Instance Segmentation at Sub-Cubic Millimeter Scale

Zudi Lin^{1†}, Donglai Wei^{1†}, Mariela D. Petkova¹, Yuelong Wu¹, Zergham Ahmed¹, Krishna Swaroop K^{2*}, Silin Zou¹, Nils Wendt^{3*}, Jonathan Boulanger-Weill¹, Xueying Wang¹, Nagaraju Dhanyasi¹, Ignacio Arganda-Carreras^{4,5,6}, Florian Engert¹, Jeff Lichtman¹, and Hanspeter Pfister¹

¹Harvard University ²NIT Karnataka ³Technical University of Munich ⁴Donostia International Physics Center (DIPC) ⁵University of the Basque Country (UPV/EHU) ⁶Ikerbasque, Basque Foundation for Science
{linzudi,donglai}@g.harvard.edu

Abstract. Segmenting 3D cell nuclei from microscopy image volumes is critical for biological and clinical analysis, enabling the study of cellular expression patterns and cell lineages. However, current datasets for *neuronal* nuclei usually contain volumes smaller than 10^{-3} mm^3 with fewer than 500 instances per volume, unable to reveal the complexity in large brain regions and restrict the investigation of neuronal structures. In this paper, we have pushed the task forward to the sub-cubic millimeter scale and curated the *NucMM* dataset with two fully annotated volumes: one 0.1 mm^3 electron microscopy (EM) volume containing nearly the entire zebrafish brain with around 170,000 nuclei; and one 0.25 mm^3 micro-CT (uCT) volume containing part of a mouse visual cortex with about 7,000 nuclei. With two imaging modalities and significantly increased volume size and instance numbers, we discover a great diversity of neuronal nuclei in appearance and density, introducing new challenges to the field. We also perform a statistical analysis to illustrate those challenges quantitatively. To tackle the challenges, we propose a novel hybrid-representation learning model that combines the merits of foreground mask, contour map, and signed distance transform to produce high-quality 3D masks. The benchmark comparisons on the NucMM dataset show that our proposed method significantly outperforms state-of-the-art nuclei segmentation approaches. Code and data are available at <https://connectomics-bazaar.github.io/proj/nucMM/index.html>.

Keywords: 3D Instance Segmentation · Nuclei · Brain · Electron Microscopy (EM) · Micro-CT (uCT) · Zebrafish · Mouse

1 Introduction

Segmenting cell nuclei from volumetric (3D) microscopy images is an essential task in studying biological systems, ranging from specific tissues [20] to entire

[†] Equally contributed.

* Works were done during internship at Harvard University.

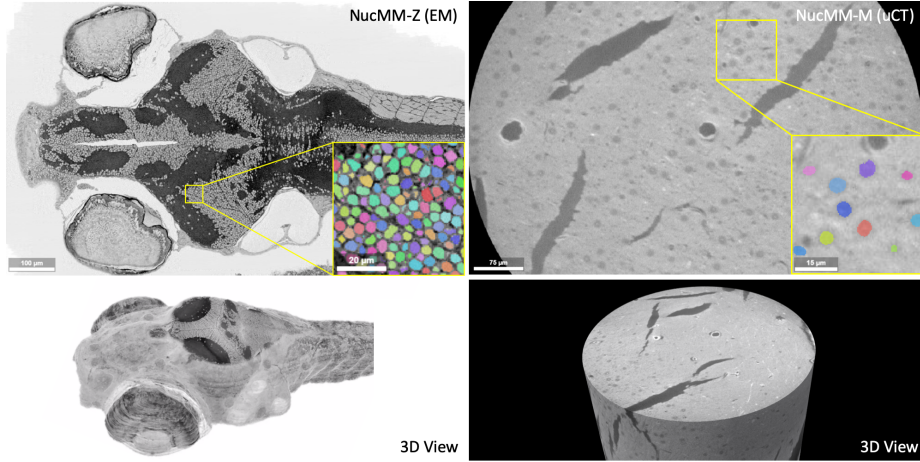


Fig. 1: **Overview of the NucMM dataset.** NucMM contains two large volumes for 3D nuclei instance segmentation, including **(left)** the NucMM-Z electron microscopy (EM) volume covering nearly a whole zebrafish brain, and **(right)** the NucMM-M micro-CT (uCT) volume from the visual cortex of a mouse.

organs [28] and developing animals [18,25]. There has been a great success in benchmarking 2D and 3D nuclei segmentation methods using datasets covering samples from various species [4,33,27,30]. However, existing datasets only have relatively small samples from brain tissues (*e.g.*, volumes from mouse and rat brains [23] smaller than 10^{-3} mm^3 , with less than 200 instances each), restricting the investigation of neuronal nuclei in a larger and more diverse scale. Besides, most of the images are collected with optical microscopy, which can not reflect the challenges in other imaging modalities widely used in studying brain tissues, including electron microscopy (EM) [14,24] and micro-CT (uCT) [9].

To address this deficiency in the field, we have curated a large-scale 3D nuclei instance segmentation dataset, **NucMM**, which is over two magnitudes larger in terms of volume size and the number of instances than previous neuronal nuclei datasets [4,23] and widely-used non-neuronal benchmark datasets [1,33]. Our NucMM consists of one EM image volume covering a nearly entire larval zebrafish brain and a uCT image volume from the visual cortex of an adult mouse, facilitating large-scale cross-tissue and cross-modality comparison. Challenges in the two volumes for automatic approaches include the high density of closely touching instances (Fig. 1, left) and low contrast between object and non-object regions (Fig. 1, right). We also perform a statistical analysis to provide a quantitative justification of the challenges in these two volumes (Fig. 2).

To tackle the challenges introduced by the large-scale NucMM dataset, we propose a new hybrid-representation model that learns foreground mask, instance contour map, and signed distance transform simultaneously with a 3D U-Net [5] architecture, which is denoted as U3D-BCD. At inference time, we

utilize a multi-target watershed segmentation algorithm that combines all three predictions to separate closely touching instances and generate high-quality instance masks. Under the average precision (AP) metric for evaluating instance segmentation approaches, we show that our U3D-BCD model significantly outperforms existing state-of-the-art approaches by **22%** on the NucMM dataset. We also perform ablation studies on the validation data to demonstrate the sensitivity of segmentation parameters.

To summarize, we have three main contributions in this paper. First, we collected and densely annotated the NucMM dataset, which is one of the largest public neuronal nuclei instance segmentation datasets to date, covering two species with two imaging modalities. Second, we propose a novel hybrid representation model, U3D-BC, to produce high-quality predictions by combining the merits of different mask representations. Third, we benchmark state-of-the-art nuclei segmentation approaches and show that our proposed model outperforms existing approaches by a large margin on the NucMM dataset.

1.1 Related Works

Nuclei Segmentation Datasets. Automatic cell nuclei segmentation is a long-lasting problem that is usually the first step in microscopy image analysis [19]. There has been a great success in benchmarking nuclei segmentation algorithms in 2D images across various types and experimental conditions [4]. However, 2D images can not completely display the structure and distribution of nuclei. Therefore, several 3D nuclei segmentation datasets have been collected, including the *Parhyale* dataset showing the histone fluorescent protein expression of *Parhyale hawaiiensis* [1,33], the BBBC050 dataset showing nuclei of mouse embryonic cells [27], and multiple 3D volumes recording human cancer cells and animal embryonic cells [30]. However, there are few datasets covering neuronal nuclei in brain tissues, and the volumes collected from brain tissue usually contain less than 200 instances per volume [4,23]. Besides, most public datasets mentioned above only have fluorescence images obtained with optical microscopy.

The NucMM dataset we curated contains two volumes from animal brain tissues and covers two imaging modalities widely used in neuroscience, including EM [14,24] and uCT [9]. In addition, our dataset is over two magnitudes larger than previous neuronal nuclei volumes [23] in terms of size and instance number.

Instance Segmentation in Microscopy Images. Instance segmentation requires assigning each pixel (voxel) not only a semantic label but also an instance index to differentiate objects that belong to the same category. The permutation-invariance of object indices makes the task challenging. For segmenting instances in microscopy images, recent learning-based approaches first train 2D or 3D convolutional neural network (CNN) models to predict an intermediate representation of the object masks such as boundary [6,22] or affinity maps [29,16]. Then techniques including watershed transform [8,35] and graph partition [15] are applied to convert the representations into object masks. Since one representation can be vulnerable to some specific kinds of errors (*e.g.*, small

Table 1: **NucMM dataset characteristics.** We collected and fully annotated a *neuronal nuclei* segmentation dataset at the sub-cubic millimeter scale. The two volumes in the dataset cover two species and two imaging modalities.

Name	Sample	Modality	Volume Size	Resolution (μm)	#Instance
NucMM-Z	Zebrafish	SEM	$397 \times 1450 \times 2000$ (0.14mm^3)	$0.48 \times 0.51 \times 0.51$	170K
NucMM-M	Mouse	Micro-CT	$700 \times 996 \times 968$ (0.25mm^3)	$0.72 \times 0.72 \times 0.72$	7K

mispredictions on the affinity or boundary map can cause significant merge errors), some approaches also employ hybrid-representation learning models that learn multiple representations and combine their information in the segmentation step [26,33,32]. However, those approaches only optimize 2D CNNs [26] or learn targets that may not be suitable for nuclei segmentation [32], which leads to unsatisfactory results for down-stream analysis. Thus we propose a 3D hybrid-representation learning model that predicts foreground mask, instance contour, and signed distance to better capture nuclei with different textures and shapes.

2 NucMM Dataset

Dataset Acquisition. The EM dataset was collected from an entire larval zebrafish brain with *serial-section* electron microscopy (SEM) [21]. The original resolution is $30\text{nm} \times 4\text{nm} \times 4\text{nm}$ for z -, y -, x -axis. We downsample the images to $480\text{nm} \times 512\text{nm} \times 512\text{nm}$ (Table 1) to make the resolution close to other imaging modalities for nuclei analysis. The micro-CT dataset was collected from layer II/III in the primary visual cortex of an adult male mouse using 3D X-ray microscopy. The images have an isotropic voxel size of 720nm^3 (Table 1).

Dataset Annotation. For annotating both volumes, we use a semi-automatic pipeline that first applies automatic algorithms to generate instance candidates and then asks neuroscientists to proofread the masks. Since EM images have high contrast and the object boundaries are well-defined (Fig. 1), we apply filtering and thresholding to get the binary foreground mask and run a watershed transform to produce the segments. For the uCT data, we iteratively enlarge the labeled set by alternating between manual correction (annotation) and automatic U-Net [22] prediction. Both volumes are finally proofread by experienced neuroscientists using VAST [2]. We also provide a binary mask for each volume to indicate the valid brain region for evaluation.

Dataset Statistics. The EM data has a size of $397 \times 1450 \times 2000$ voxels, equivalent to a physical size of 0.14mm^3 ; the uCT data has $700 \times 996 \times 968$ voxels, equivalent to 0.25mm^3 (Table 1). Although the zebrafish volume is physically smaller, it contains significantly more objects than the mouse volume, showing the distinction between brain structures. We show the distribution of nuclei size (Fig. 2a) and the nearest-neighbor distance between nuclei centers (Fig. 2b). The density plots illustrate that objects in the zebrafish volume are smaller and more densely packed, which poses challenges in separating closely touching nuclei.

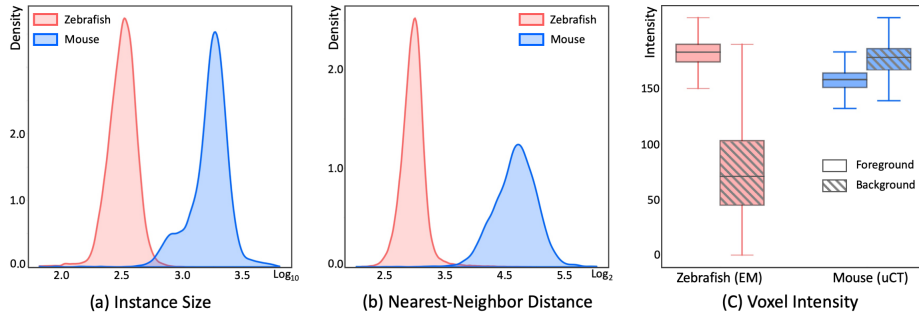


Fig. 2: **NucMM dataset statistics.** We show the distribution of (a) instance size (in terms of number of voxels) and (b) the distance between adjacent nuclei centers. The density plots are normalized by the total number of instances in each volume. We also show (c) the voxel intensity distribution in object (foreground) and non-object (background) regions for both volumes.

We further show the voxel intensity distribution of object and non-object regions (Fig. 2c). The boxplots demonstrate that objects in the EM volume are better separated from the background than the uCT volume. To quantify the separation, we calculate the *Kullback–Leibler* (KL) divergence between the foreground and background intensity distributions in the two volumes. The results are $D_{KL} = 3.43$ and 1.33 for the zebrafish and mouse data, respectively, showing the foreground-background contrast is significantly lower in the uCT data.

Dataset Splits & Evaluation Metric. We split each volume into 5%, 5%, and 90% parts for training, validation, and testing. The limited training data makes the task more challenging, but it is also closer to the realistic annotation budget when neuroscientists handle newly collected data. Besides, to avoid sampling data in a local region without enough diversity, we follow previous practice [13] and sample 27 small chunks of size $64 \times 64 \times 64$ voxels from the zebrafish volume for training. Since the nuclei in the mouse uCT volume are sparser (Fig. 1 and Fig. 2b), we sample 4 chunks of size $192 \times 192 \times 192$ voxels for training.

For evaluation, we use average precision (AP), a standard metric in assessing instance segmentation methods [7, 17]. Specifically, we use the code optimized for large-scale 3D image volumes [32] to facilitate efficient evaluation.

3 Method

3.1 Hybrid-Representation Learning

Recent 3D instance segmentation methods for microscopy images, including Cellpose [26], StarDist [33] and U3D-BC [32], all use a single model to learn multiple mask representations simultaneously. Specifically, Cellpose [26] regresses the horizontal and vertical spatial gradients of the instances; StarDist [33] learns object probability and the star-convex distance within the masks; U3D-BC [32] learns

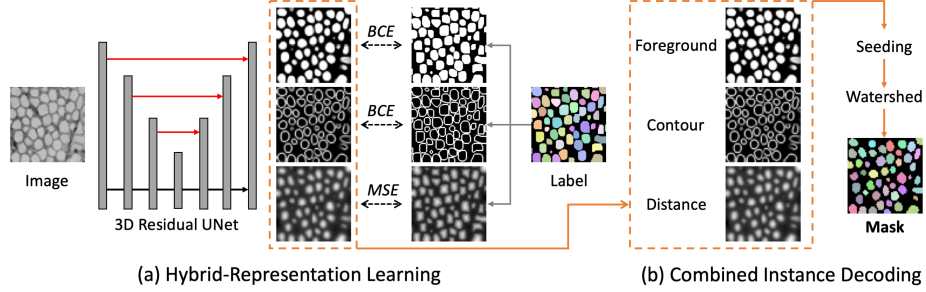


Fig. 3: **Hybrid-representation learning model.** (a) Our U3D-BCD model learns a set of hybrid representations simultaneously, including foreground mask, instance contour, and signed distance transform map calculated from the segmentation. (b) The representations are combined in seeding and watershed transform to produce high-quality segmentation masks.

the foreground mask with the instance contour map. By analyzing the representations used in those models, we notice that all the targets emphasize the learning of object masks (foreground), but the structure of the *background* of the segmentation map is less utilized. That is, pixels close to and far away from the object masks are treated equitably.

Therefore based on the U3D-BC model that predicts both the binary foreground mask (B) and instance contour map (C), we develop a **U3D-BCD** model that in addition predicts a signed Euclidean distance map (Fig. 3a). Let x_i denote a pixel in the image, we have:

$$f(x_i) = \begin{cases} +\text{dist}(x_i, B)/\alpha, & \text{if } x \in F. \\ -\text{dist}(x_i, F)/\beta, & \text{if } x \in B. \end{cases} \quad (1)$$

where F and B denote the foreground and background masks, respectively. The scaling parameters α and β are applied to control the range of the distance. Compared with the U3D-BC baseline, the signed distance map is more informative in capturing the shape information of masks. In comparison with all discussed approaches [26, 33, 32], the signed distance map also model the landscape of background regions. In implementation, we apply a *tanh* activation to restrict the range to $(-1, 1)$ and directly regress the target with a 3D U-Net [5]. A similar learning target has been used for semantic segmentation of synapses [12], but it has not been integrated into a multi-task learning model nor has it been explored for 3D instance segmentation. Specifically, the loss we optimize is

$$\mathcal{L}_{bcd} = h(\sigma(y_1), y_b) + h(\sigma(y_2), y_c) + g(\phi(y_3), y_d) \quad (2)$$

where y_i ($i = 1, 2, 3$) denote the three output channels, while σ and ϕ denote the sigmoid and tanh function. The foreground (y_b) and contour (y_c) maps are learned by optimizing the binary cross-entropy (BCE) loss (h), while the signed distance map (y_d) is learned with the mean-squared-error (MSE) loss (g).

3.2 Instance Decoding

In this part, we describe how to combine all three model predictions in the U3D-BCD model to generate the segmentation masks. We first threshold predictions to locate *seeds* (or markers) with high foreground probability and distance value but low contour probability, which points to object centers. We then apply the marker-controlled watershed transform algorithm (in the *scikit-image* library [31]) using the seeds and the predicted distance map to produce masks (Fig. 3b). There are two advantages over the U3D-BC [32] model, which also uses marker-controlled watershed transform for decoding. First, we make use of the consistency among the three representations to locate the seeds, which is more robust than U3D-BC that uses two predictions. Furthermore, we use the smooth signed distance map in watershed decoding, which can better capture instance structure than the foreground probability map in U3D-BC. We also show the sensitivity of decoding parameters in the experiments.

3.3 Implementation

We use a customized 3D U-Net model that substitutes the convolutional layers at each stage with residual blocks [10]. We also change the concatenation operation to addition to save memory. Since recent work [34] has shown that ADAM-like adaptive optimization algorithms do not generalize as well as stochastic gradient descent (SGD) [3], we optimize our models with SGD and adopt *cosine learning rate decay* [11]. We set the scaling parameters α and β of the signed distance map to 8 and 50, respectively, without tweaking. We follow the open-source code of U3D-BC¹ and apply data augmentations including brightness, flip, rotation, elastic transform, and missing parts augmentations to U3D-BC [32] and our U3D-BCD. We also use the official implementation of StarDist² and Cellpose³.

4 Experiments

4.1 Methods in Comparison

We benchmark state-of-the-art microscopy image segmentation models including Cellpose [26], StarDist [33] and U3D-BC [32] based on their public implementations. Specifically, the Cellpose model is trained on the 2D xy , yz , and xz planes from the image volumes, and the predicted spatial gradients are averaged to generate a 3D gradient before segmentation. For StarDist, we calculate the optimal number of rays on the training data, which are 96 and 64 for the zebrafish and mouse volumes, respectively. For U3D-BC, we use the default 1.0 weight ratio between the losses of the foreground and contour map. Models are trained on a machine with four Tesla V100 GPUs.

¹ https://github.com/zudi-lin/pytorch_connectomics

² <https://github.com/stardist/stardist>

³ <https://github.com/MouseLand/cellpose>

Table 2: **Benchmark results on the NucMM dataset.** We compare state-of-the-art methods on the NucMM dataset using AP score. **Bold** and underlined numbers denote the 1st and 2nd scores, respectively. Our U3D-BCD model significantly improves the performance of previously state-of-the-art approaches.

Method	NucMM-Z			NucMM-M			Overall
	AP-50	AP-75	Mean	AP-50	AP-75	Mean	
Cellpose [26]	0.796	0.342	0.569	0.463	0.002	0.233	0.401
StarDist [33]	<u>0.912</u>	0.328	0.620	0.306	0.004	0.155	0.388
U3D-BC [32]	0.782	<u>0.556</u>	<u>0.670</u>	0.645	<u>0.210</u>	<u>0.428</u>	<u>0.549</u>
U3D-BCD (Ours)	0.978	0.809	0.894	<u>0.638</u>	0.250	0.444	0.669

4.2 Benchmark Results on the NucMM Dataset

After choosing hyper-parameter on the validation sets, we run predictions on the 90% test data in each volume and evaluate the performance. Specifically, we show the AP scores at intersection-over-union (IoU) thresholds of both 0.5 (AP-50) and 0.75 (AP-75), as well as their average (Table 2). The overall score is averaged over two NucMM volumes. The results show that our U3D-BCD model significantly outperforms previous state-of-the-art models by relatively **22%** in overall performance. Besides, our method ranks 1st in 6 out of 7 scores, showing its robustness in handling different challenges. We argue that Cellpose [26] trains 2D models to estimate 3D spatial gradient, which can be ineffective for challenging 3D cases. The other two models use 3D models, but the representations StarDist [33] uses overlooks the background in the segmentation mask, while U3D-BC [32] overlooks both foreground and background structures. Although conceptually straightforward, introducing the signed distance on top of the U3D-BC baseline gives a notable performance boost.

4.3 Sensitivity of the decoding parameters

We also show the sensitivity of the decoding hyper-parameters of our U3D-BCD model. Specifically, there are 5 thresholds for segmentation: 3 values are the thresholds of foreground probability (τ_1), contour probability (τ_2), and distance value (τ_3) to decide the seeds. The other 2 values are the thresholds of foreground probability (τ_4) and distance value (τ_5) to decide the valid foreground regions.

The validation results show that the final segmentation performance is not sensitive to the τ_1 and τ_2 in deciding seeds. When fixing $\tau_4 = 0.2$ and $\tau_5 = 0.0$ changing the foreground probability τ_1 from 0.4 to 0.8 only changes the AP-50 score between 0.943 and 0.946. While changing the contour probability τ_2 from 0.05 to 0.30 only changes the scores from 0.937 to 0.946. However, when changing τ_4 within the range of 0.1 to 0.4, the score changes from 0.872 to 0.946. Those results suggest that the signed distance transform map contains important information about the object structures and performs an important role in generating high-quality segmentation masks.

5 Conclusion

In this paper, we introduce the large-scale NucMM dataset for 3D neuronal nuclei segmentation in two imaging modalities and analysis the challenges quantitatively. We also propose a simple yet effective model that significantly outperforms existing approaches. We expect the densely annotated dataset can inspire various applications beyond its original task, *e.g.*, feature pre-training, shape analysis, and benchmarking active learning and domain adaptation methods.

Acknowledgments

This work has been partially supported by NSF award IIS-1835231 and NIH award U19NS104653. We thank Daniel Franco-Barranco for setting up the challenge using NucMM. M.D.P. would like to acknowledge the support of Howard Hughes Medical Institute International Predoctoral Student Research Fellowship. I.A-C would like to acknowledge the support of the Beca Leonardo a Investigadores y Creadores Culturales 2020 de la Fundación BBVA.

References

1. Alves, F., Enjolras, C., Averof, M.: Live imaging reveals the progenitors and cell dynamics of limb regeneration. *Elife* (2016)
2. Berger, D.R., Seung, H.S., Lichtman, J.W.: Vast (volume annotation and segmentation tool): efficient manual and semi-automatic labeling of large 3D image stacks. *Frontiers in neural circuits* (2018)
3. Bottou, L.: Stochastic gradient learning in neural networks. *Proceedings of NeuroNimes* (1991)
4. Caicedo, J.C., Goodman, A., Karhohs, K.W., Cimini, B.A., Ackerman, J., Haghighi, M., Heng, C., Becker, T., Doan, M., McQuin, C., et al.: Nucleus segmentation across imaging experiments: the 2018 data science bowl. *Nature methods* (2019)
5. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3d unet: learning dense volumetric segmentation from sparse annotation. In: *MICCAI* (2016)
6. Ciresan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: *NeurIPS* (2012)
7. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B.: The cityscapes dataset for semantic urban scene understanding. In: *CVPR* (2016)
8. Cousty, J., Bertrand, G., Najman, L., Couprie, M.: Watershed cuts: Minimum spanning forests and the drop of water principle. *TPAMI* (2008)
9. Dyer, E.L., Roncal, W.G., Prasad, J.A., Fernandes, H.L., Gürsoy, D., De Andrade, V., Fezzaa, K., Xiao, X., Vogelstein, J.T., Jacobsen, C., et al.: Quantifying mesoscale neuroanatomy using x-ray microtomography. *eneuro* (2017)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *CVPR* (2016)
11. He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J., Li, M.: Bag of tricks for image classification with convolutional neural networks. In: *CVPR* (2019)

12. Heinrich, L., Funke, J., Pape, C., Nunez-Iglesias, J., Saalfeld, S.: Synaptic cleft segmentation in non-isotropic volume electron microscopy of the complete drosophila brain. In: MICCAI (2018)
13. Januszewski, M., Kornfeld, J., Li, P.H., Pope, A., Blakely, T., Lindsey, L., Maitin-Shepard, J., Tyka, M., Denk, W., Jain, V.: High-precision automated reconstruction of neurons with flood-filling networks. *Nature Methods* (2018)
14. Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T.R., et al.: Saturated reconstruction of a volume of neocortex. *Cell* (2015)
15. Krasowski, N., Beier, T., Knott, G., Köthe, U., Hamprecht, F.A., Kreshuk, A.: Neuron segmentation with high-level biological priors. *TMI* (2017)
16. Lee, K., Zung, J., Li, P., Jain, V., Seung, H.S.: Superhuman accuracy on the SNEMI3D connectomics challenge. *arXiv:1706.00120* (2017)
17. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: *ECCV* (2014)
18. Lou, X., Kang, M., Xenopoulos, P., Munoz-Descalzo, S., Hadjantonakis, A.K.: A rapid and efficient 2d/3d nuclear segmentation method for analysis of early mouse embryo and stem cell image data. *Stem cell reports* (2014)
19. Meijering, E.: Cell segmentation: 50 years down the road. *Signal Processing Magazine* (2012)
20. Nhu, H.T.T., Drigo, R.A.E., Berggren, P.O., Boudier, T.: A novel toolbox to investigate tissue spatial organization applied to the study of the islets of langerhans. *Scientific reports* (2017)
21. Petkova, M.: Correlative Light and Electron Microscopy in an Intact Larval Zebrafish. Ph.D. thesis (2020)
22. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *MICCAI* (2015)
23. Ruszczycki, B., Pels, K.K., Walczak, A., Zamłyńska, K., Such, M., Szczepankiewicz, A.A., Hall, M.H., Magalska, A., Magnowska, M., Wolny, A., et al.: Three-dimensional segmentation and reconstruction of neuronal nuclei in confocal microscopic images. *Frontiers in neuroanatomy* (2019)
24. Shapson-Coe, A., Januszewski, M., Berger, D.R., Pope, A., Wu, Y., Blakely, T., Schalek, R.L., Li, P., Wang, S., Maitin-Shepard, J., et al.: A connectomic study of a petascale fragment of human cerebral cortex. *bioRxiv* (2021)
25. Stegmaier, J., Amat, F., Lemon, W.C., McDole, K., Wan, Y., Teodoro, G., Mikut, R., Keller, P.J.: Real-time three-dimensional cell segmentation in large-scale microscopy data of developing embryos. *Developmental cell* (2016)
26. Stringer, C., Wang, T., Michaelos, M., Pachitariu, M.: Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods* (2021)
27. Tokuoka, Y., Yamada, T.G., Mashiko, D., Ikeda, Z., Hiroi, N.F., Kobayashi, T.J., Yamagata, K., Funahashi, A.: 3d convolutional neural networks-based segmentation to acquire quantitative criteria of the nucleus during mouse embryogenesis. *NPJ systems biology and applications* (2020)
28. Toyoshima, Y., Tokunaga, T., Hirose, O., Kanamori, M., Teramoto, T., Jang, M.S., Kuge, S., Ishihara, T., Yoshida, R., Iino, Y.: Accurate automatic detection of densely distributed cell nuclei in 3d space. *PLoS computational biology* (2016)
29. Turaga, S.C., Briggman, K.L., Helmstaedter, M., Denk, W., Seung, H.S.: Maximin affinity learning of image segmentation. In: *NeurIPS* (2009)
30. Ulman, V., Maška, M., Magnusson, K.E., Ronneberger, O., Haubold, C., Harder, N., Matula, P., Matula, P., Svoboda, D., Radojevic, M., et al.: An objective comparison of cell-tracking algorithms. *Nature methods* (2017)

31. van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., the scikit-image contributors: scikit-image: image processing in Python. *PeerJ* (2014)
32. Wei, D., Lin, Z., Franco-Barranco, D., Wendt, N., et al.: MitoEM dataset: Large-scale 3D mitochondria instance segmentation from EM images. In: *MICCAI* (2020)
33. Weigert, M., Schmidt, U., Haase, R., Sugawara, K., Myers, G.: Star-convex polyhedra for 3d object detection and segmentation in microscopy. In: *WACV* (2020)
34. Zhou, P., Feng, J., Ma, C., Xiong, C., HOI, S., et al.: Towards theoretically understanding why sgd generalizes better than adam in deep learning. *arXiv preprint arXiv:2010.05627* (2020)
35. Zlateski, A., Seung, H.S.: Image segmentation by size-dependent single linkage clustering of a watershed basin graph. *arXiv:1505.00249* (2015)