Supporting Information

Pelt and Sethian 10.1073/pnas.1715832114

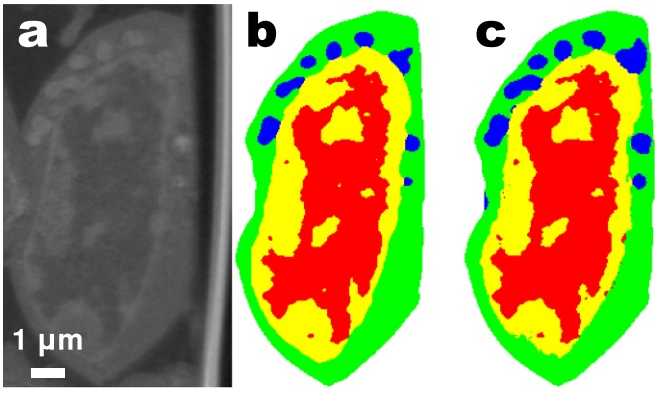

Fig. S1. (*A*–*C*) A tomographic slice of a lympoblastoid cell (*A*), with the corresponding manual segmentation (*B*) and output of an MS-D network with 100 layers (*C*).

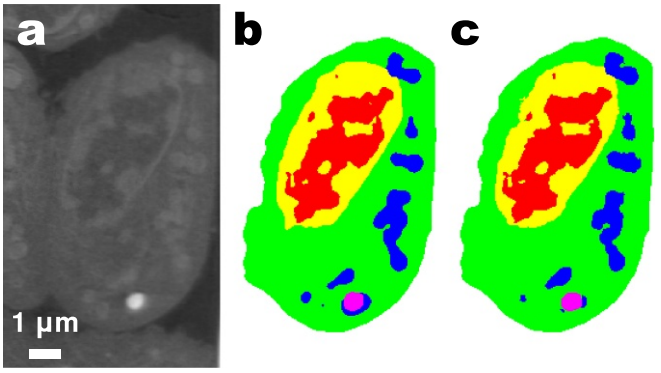

Fig. S2. (*A*–*C*) A tomographic slice of a lympoblastoid cell (*A*), with the corresponding manual segmentation (*B*) and output of an MS-D network with 100 layers (*C*).

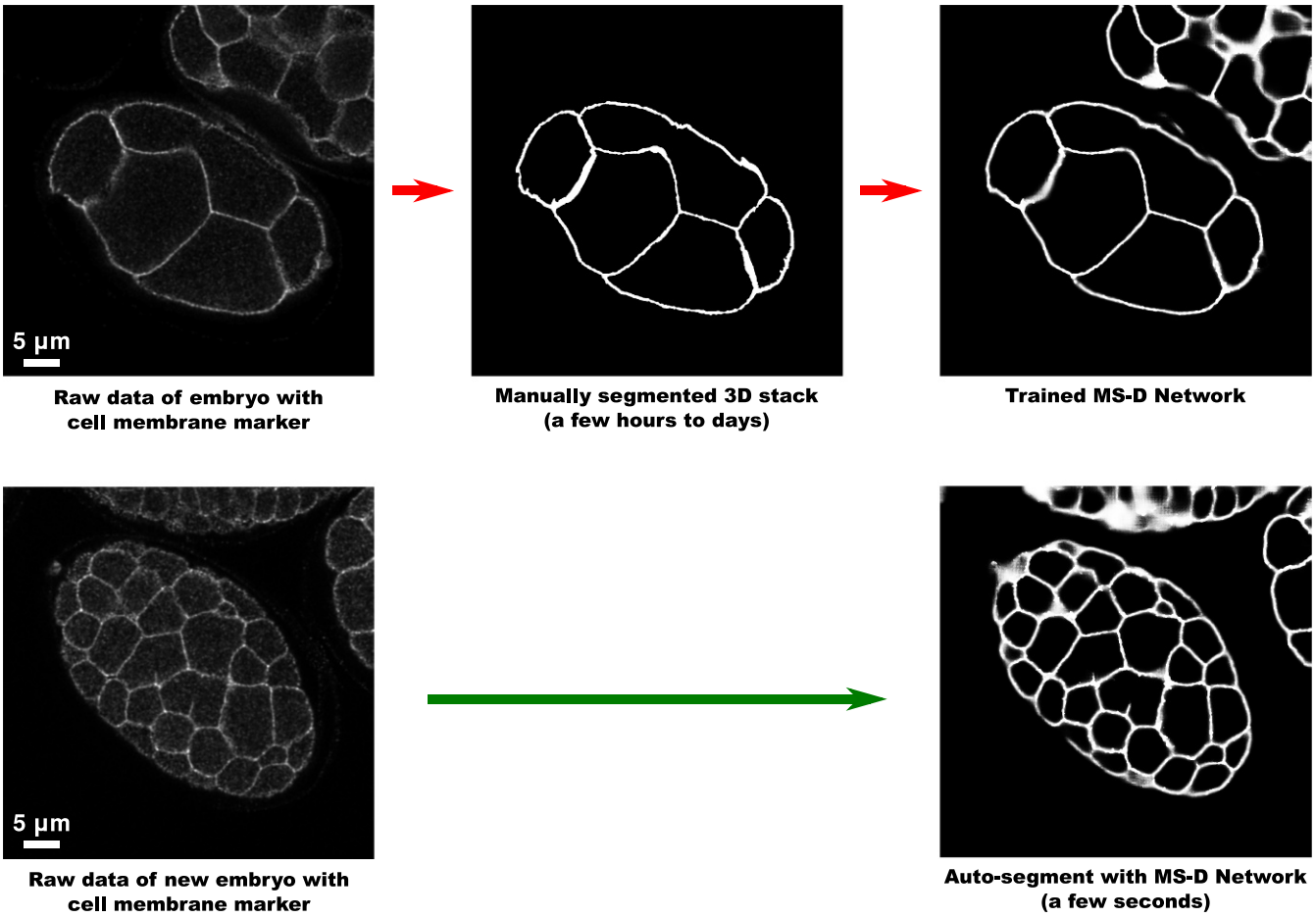

Fig. S3. Example of how an MS-D network can improve the analysis of *Caenorhabditis elegans* worm embryos by eliminating time-consuming manual segmentation of the cell membranes. Here, we used an MS-D network with 100 layers and trained using a single manually segmented stack of 190 images from a single *C. elegans* embryo that was engineered to express a GFP-tagged cell membrane protein. (*Top Row*) To generate the training material for the machine-learning neural network, a stack of 190 images from a seven-cell embryo was first thresholded. The speckles were then removed from each image, and the gaps in the membranes were filled in manually to reflect the actual cell membrane structure. These 190 manually segmented images were then trained to generate an algorithm for labeling cell boundaries. (*Bottom Row*) The algorithm generated by machine learning was then directly applied to the raw images of an 86-cell *C. elegans* embryo to label boundaries for all of the cells within the embryo. *C. elegans* embryo contributions are attributed to S. Uzawa, Q. Bian, and B. J. Meyer, Howard Hughes Medical Institute and Department of Molecular and Cell Biology at University of California, Berkeley.

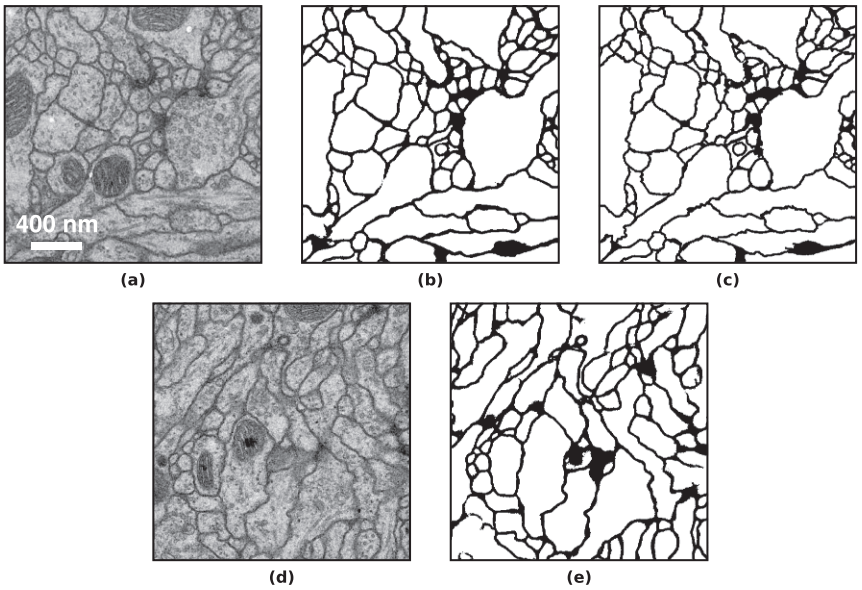

Fig. S4. Example of applying an MS-D network with $w = 1$ and $d = 100$ to the "International Symposium on Biomedical Imaging (ISBI) challenge: Segmentation of neuronal structures in EM stacks" (1). Input images consist of 512 × 512-pixel serial section transmission electron microscopy (ssTEM) images of the *Drosophila* first instar larva ventral nerve cord (VNC). The goal is to segment neural structures in each image. The challenge dataset consists of 30 training images for which a manual labeling is provided and 30 testing images for which the manual labeling is withheld. For training the MS-D network, the 30 available images were augmented by rotation, reflection, and elastic deformation, and training was stopped after no improvement was observed in the global accuracy for the original, not augmented, images. In *A*, an input image from the training set is shown, with the corresponding output of the MS-D network in *B*, and manual labeling in *C*. An input image from the test set is shown in *D*, with the corresponding MS-D network output shown in *E*. For more information about the challenge, see ref. 1.

1. Arganda-Carreras I, et al. (2015) Crowdsourcing the creation of image segmentation algorithms for connectomics. *Front Neuroanat* 9:142.

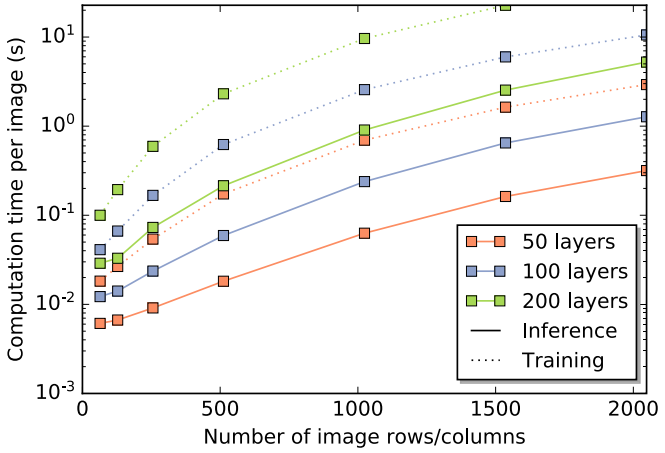

Fig. S5. The computation time of processing a single image with an MS-D network as a function of the number of rows and columns, for the application shown in Fig. 8. Results are shown for both inference (i.e., a forward pass) and training (i.e., backpropagation and gradient computation), for various numbers of layers. Computation times were measured by taking the average time of 500 computations, using a single GTX 1080 GPU and a batch size of one. Note that, in our implementation, the required computation time for larger batches scales linearly with the number of images.