

---

# AUTOMATED SEGMENTATION OF FIB-SEM NUCLEUS IMAGES

---

**Niraj Gupta\***

Department of Computer Science  
University of Colorado, Boulder  
Boulder, CO 80309, USA  
niraj.gupta@colorado.edu

**Eric J. Roberts\***

Molecular Biophysics & Integrated Bioimaging Division  
Center for Advanced Mathematics for Energy Research Applications  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720, USA  
EJRoberts@lbl.gov

**Song Pang**

Janelia Research Campus  
Howard Hughes Medical Institute  
Ashburn, VA 20147, USA.  
pangs@janelia.hhmi.org

**C. Shan Xu**

Janelia Research Campus  
Howard Hughes Medical Institute  
Ashburn, VA 20147, USA.  
xuc@janelia.hhmi.org

**Harald F. Hess**

Janelia Research Campus  
Howard Hughes Medical Institute  
Ashburn, VA 20147, USA.  
hessh@janelia.hhmi.org

**Petrus H. Zwart**

Molecular Biophysics & Integrated Bioimaging Division  
Center for Advanced Mathematics for Energy Research Applications  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720, USA  
PHZwart@lbl.gov

**Vignesh Kasinath†**

Assistant Professor  
Department of Biochemistry  
University of Colorado, Boulder  
Boulder, CO 80309, USA  
vignesh@colorado.edu

July 17, 2022

## ABSTRACT

Let's add the abstract last

## 1 Introduction

Focused ion beam scanning electron microscopy (FIB-SEM) is an automated imaging technique capable of probing structure on the nanometer scale. FIB-SEM consists of dual beam instrumentation in which a scanning electron microscope works in combination with high-energy ion beams that mill away, *in situ*, a resin-encased specimen. In this setup, the microscope directly observes the newly revealed ultrathin layer of the sample as it is milled away. The serial images are then collected and aligned to produce a 3d volume image stack. This led to unprecedented *z*-axis resolution of fewer than ten nanometers, resulting in isotropic (i.e. *z*-axis sectional thickness is equal to *x*- and *y*-axis pixel size), full image resolutions and user-specified site-specific selections of 4nm<sup>3</sup>-sized voxels, achieved in a pioneering study of aldehyde-fixed mouse brain tissue images [1]. This represents an order of magnitude upgrade in *z*-axis resolution over the best current SEM-only techniques, which alternate between imaging and cutting away the surface with diamond-tipped knives [2, 3]. While these and similar cutting strategies were significant advancements for biological imaging – this represented the first instance of a fully automated, *in situ* volumetric EM approach performed by cutting material layers off [4] – these methods lose consistency when cutting less than 25 nm deep [5, 6], thus limiting both resolution and the acquisition volume. In replacing the diamond knife with the ion beam, FIB-SEM approaches, but does not reach, the state-of-the-art resolution for nanometer-scale 3D imaging of ~1 nm-sized voxels, achieved by transmission electron microscopy (TEM) tomographic in which electrons are transmitted through the medium or sample. These transmission techniques are typically limited to sections of material only ~ 200-500 nm deep. FIB-SEM can

---

\*These authors contributed equally to this work

†Corresponding author

instead run indefinitely, allowing for far greater imaging depth and unparalleled levels of automation that is constrained only by the availability of material, time, and operational costs.

Before being applied in biological settings, FIB-SEM was first proposed and implemented in soft matter physics and has long proven invaluable to material scientists, leading to effective analysis and classification of micro structures in porous media [7, 8, 9, 10], characterization of nanopores in coal and other reservoir rocks [11, 12], the reconstruction of polymer films [13, 14], and better optimization and design of controlled drug release coatings [15]. However, it would take some time for FIB-SEM imaging to be used in biological settings. Only since 2004 have FIB-SEM dual beam setups been explored for biological imaging tasks, beginning with a FIB used as a cutting device for exposing tissue and gland cells *n in situ* to subsequent SEM imaging [16]. Though imaging resolution was limited to relatively shallow depths on the scale of dozens of micrometers, this was a major step forward for FIB-SEM technology and marked the beginning of modern volumetric microscopy [17]. More recently, research developments in FIB-SEM methodology have matured the technique into an essential tool in the study of cellular biology [18, 19, 20, 17, 16, 21, 22]. These developmental breakthroughs include correcting for anisotropic data slicing with optical flow interpolation [23], accelerating image acquisition times via adaptive sampling that leverages multiple low-dose, quicker scans [24], software improvements incorporating backscattered electron detectors with positive sampling bias xu2017enhanced, and considerable improvements in hardware and reliability of continuous, long-term operational sessions, upwards of several months xu2017enhanced, hayworth2015ultrastructurally.

In total, the boon in FIB-SEM imaging improvements has led to a vast increase in the amount of high quality cellular data to analyze. This is perhaps best exemplified in the recent open-source availability on the web repository “OpenOrganelle [25]” of numerous whole cell and tissue 3D images of 4 nanom isotropic voxel resolution, with some volumes measuring greater than 100,000  $\mu\text{m}^3$  in size [26]. While FIB-SEM requires very little in terms of post-processing and alignment, the sheer amount of data produced is certainly a bottleneck as the ever-increasing need for manual curation of expert annotations and analyses is only exacerbated. Thus, any steps toward automation in image analysis pipelines are crucial in furthering the understanding of cellular structures and subcellular components [27].

In this study, we apply machine learning tools for the automated delineation and segmentation of intracellular structures in FIB-SEM-acquired image volumes of *Caenorhabditis elegans* (*C.elegans*) reproductive eggs. The data is considerably large and detailed, with a resolution of  $4 \times 4 \times 4 \text{ nm}^3$  for volumes measuring upwards of  $11,250 \mu\text{/meter}^3$ . Add a bit more here. Will be related to the Rekated Works section below.

The remainder of this manuscript is organized as follows: Sect. 2 discusses advances in machine learning work applied to other intracellular FIB-SEM data; Sect. 3 describes the *C. elegans* sample preparation and imaging process; Sect. 4 details the data pre-processing, neural networks, training parameters, and evaluation metrics used for analysis; Sect. 5 walks the reader through the network training and image segmentation workflow; Sect. 6 presents machine learning-based segmentation results; and Sect. 7 offers a discussion on future works and the scope of the manuscript.

## 2 Related Works

Recent machine learning developments have rapidly increased the speed at which microscopy data is analyzed and segmented, all while offering researchers tools that remain significantly more tolerant to noise than traditional computer vision techniques [28]

1. FIB-SEM technique
2. Need for machine learning/automated segmentation to combat high amounts of data
3. Current state of machine learning performance on similar data

### 2.1 Related Works

## 3 The Data

The cells used in this study originate from *Caenorhabditis elegans*, more commonly known as a nematode or *C. elegans*. *C. elegans* is generally seen as an exemplary and prototypical organism in the investigation of developmental biology [29, 30]. *C. elegans* achieved stardom as the first multicellular organism to have its entire genome sequenced [31, 32]. Studies of *C.elegans* have helped behavioral scientists map the neural circuitry that controls touch-induced locomotion [33], deduce the functions of certain touch circuitry neurons [33], investigate clues related to the evolutionary development of the circadian clock [34], and discover nested neurological dynamics/activity patterns that govern a behavioral hierarchy of motor actions across multiple time scales [35]. *C. elegans* remains the only organism to have the *entirety* of its nervous system, also known as the connectome, mapped out [36, 37, 38]

### 3.1 FIBSEM Sample Preparation

One Durcupan-embedded *C. elegans* gonad egg sample was first mounted to the top of a 1 mm copper post which was in contact with the metal-stained sample for better charge dissipation, as previously described in [39]. The vertical sample post was first trimmed to a small block of  $95 \times 80 \times 150 \mu\text{m}^3$  containing two Regions of Interest (ROI 1-2) from top to bottom. The sample block has a width perpendicular to the ion beam, and a depth in the direction of the ion beam. After FIB-SEM imaging ROI1 and ROI2, the sample was then trimmed to the second block of  $80 \times 80 \times 100 \mu\text{m}^3$  containing ROI3. The trimming was guided by X-ray tomography data obtained by a Zeiss Versa XRM-510 and optical inspection under a microtome. Thin layers of conductive material of 10-nm gold followed by 100-nm carbon were coated on the trimmed samples using a Gatan 681 High-Resolution Ion Beam Coater. The coating parameters were 6 keV, 200 nA on both argon gas plasma sources, 10 rpm sample rotation with 45-degree tilt.

### 3.2 FIB-SEM 3D large volume imaging

One FIB-SEM prepared *C. elegans* gonad egg sample was imaged sequentially by a customized Zeiss FIB-SEM system (Germini 500) previously described in [39], [40], and [26]. The block face of each ROI was imaged by a 250 pA electron beam with 0.9 keV landing energy at 200 kHz scanning rate. The  $x-y$  pixel resolution was set at 4 nm. A subsequently applied focused Ga+ beam of 15 nA at 30 keV strafed across the top surface and ablated away 4 nm of the surface. The newly exposed surface was then imaged again. The ablation-imaging cycle continued about once every 75 seconds for 9 and 6 days to complete FIB-SEM imaging ROI1 and ROI2, respectively. Such cycle extended to once every 135 seconds for 14 days to image ROI3. The acquired image stack formed a raw imaged volume, followed by post-processing of image registration and alignment via local feature matching using a Scale Invariant Feature Transform (SIFT)-based algorithm [41, 42, 43]. The aligned stack consists of final isotropic volumes of  $10 \times 20 \times 30 \mu\text{m}^3$ ,  $10 \times 20 \times 20 \mu\text{m}^3$ , and  $25 \times 15 \times 30 \mu\text{m}^3$  for ROI1, ROI2, and ROI3, respectively. The voxel size of  $4 \times 4 \times 4 \text{ nm}^3$  was maintained for each sample throughout entire volumes, which can be viewed in any arbitrary orientations.

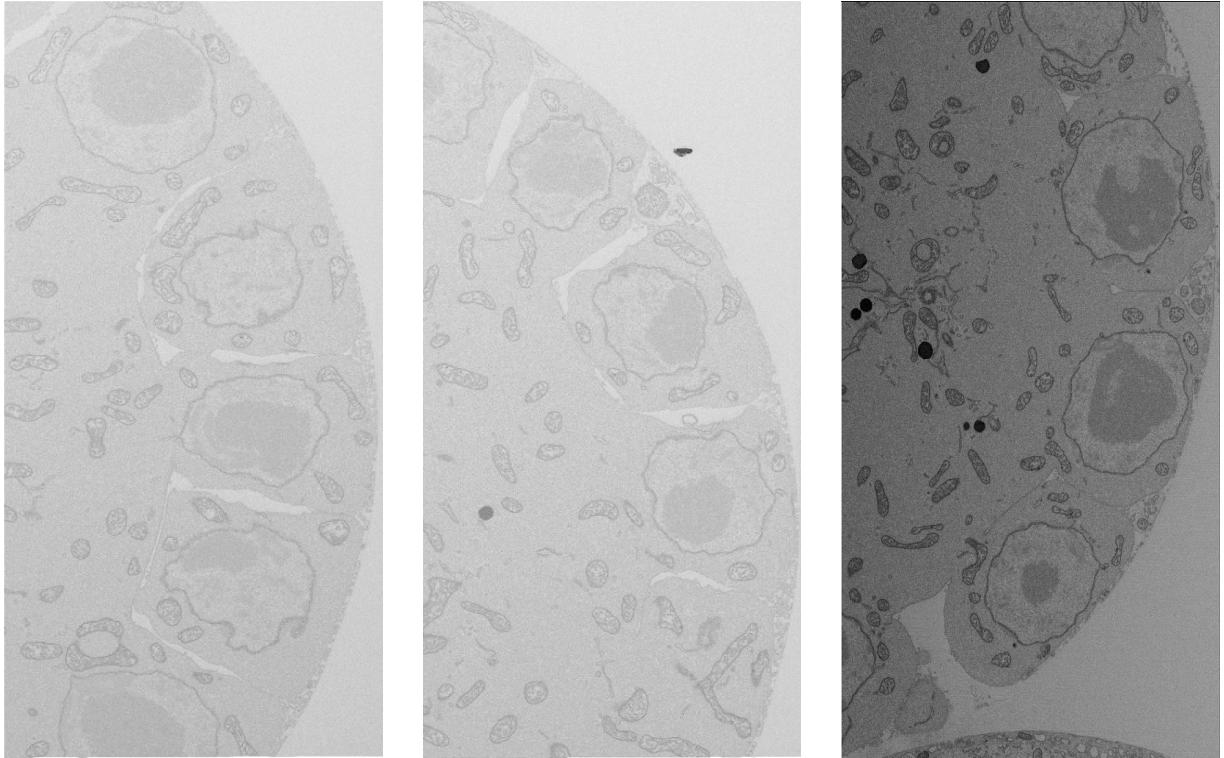


Figure 1: Individual FIB-SEM cross sectional slices of *Caenorhabditis elegans*. Pictured are (a) ROI1, (b) ROI2, and (c) ROI3 regions of interest.

## 4 Methods

### 4.1 Data Preprocessing

The images generated for the FIBSEM images are in TIFF format. ROI1 has 6801 color-scale images that are  $5000 \times 2500 \times 3$ , ROI2 has 5000 images of the same size/dimensionality, and ROI3 has 8837 images that are  $3750 \times 6250 \times 3$ . The nuclei are extracted sequentially from each tiff file using traditional computer vision techniques. Because the nucleus edges are not continuous, contour detection cannot be applied directly to the raw images. Instead, we perform image pre-processing by first minimizing image noise using Gaussian blur with an  $11 \times 11$  kernel. Gaussian blur enhances image structure by smoothing out pixel intensities [44], as pixel intensity and brightness are not uniform across the entirety of an image. Second, we perform Gaussian adaptive thresholding to help alleviate remaining inhomogeneous pixel intensities. Thresholding helps extract the margins of nuclei by producing a binary image of the nuclei outline. To fill the gaps in the nuclei boundaries, we first create a  $7 \times 7$  elliptical kernel, which is then used to calculate the morphological gradient [45, 46], which highlights stark contrasts in neighboring pixel intensities to form object outlines using the difference between the dilation and erosion morphological operators [47, 48]. The complete outlines will be fully recognized once the broken edge gaps have been filled. To accomplish this, the approximation method is set to chain approximation, and the retrieval method is set to RETR\_TREE. The chain approximation method used here stores all the boundary points of the contour [49, 50]. The tree retrieval method retrieves all of the contours by traversing over the ‘chains’ of boundary pixels, then reconstructs a full hierarchy of nested contours. These are then filtered to recover elliptical outlines within a specific size range. After filtering, the centroid is used to cluster the nuclei at the same location for any observed contour. Throughout the process of clustering nuclei based on position, a mapping of nuclei position and ID is retained. All of the nuclei collected are stored as JPEG files with a constant size of  $1700 \times 1700$  pixels.

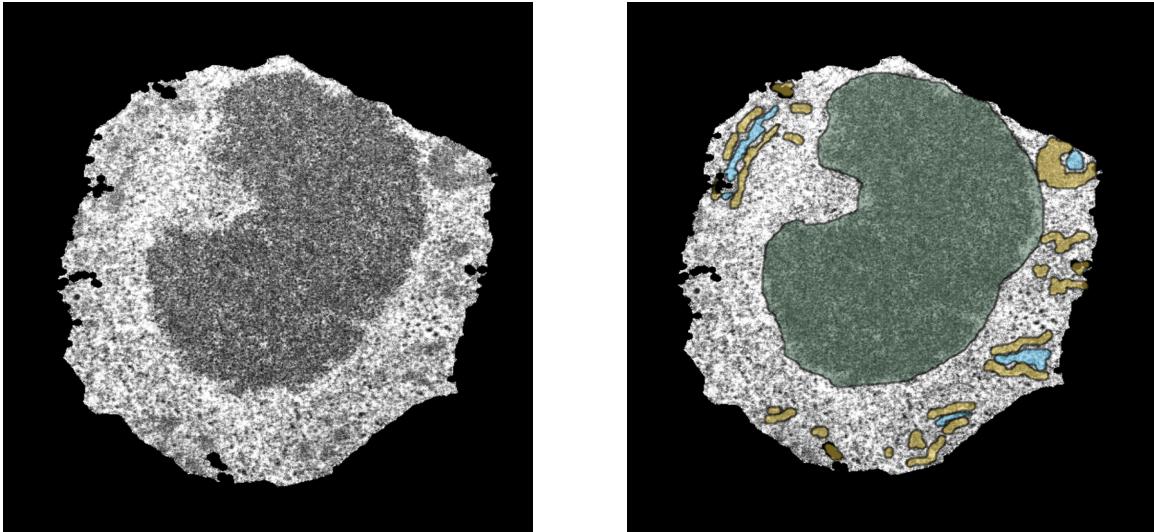


Figure 2: (a) Image of an extracted nuclei from ROI1 and (b) corresponding annotations. Here, green represents nucleolus, blue represents synaptonemal complex, and yellow represents the chromosome annotations.

Figure 2 shows the manual annotations that are created for one of the ROI1 images. The annotations are created using apeir web application (CITE apeir). All the nucleolus are labeled as 1 (green), chromosome as 2 (yellow), and the synaptonemal complex as 3 (blue) with the background as 0 for every image. For the training data, 260 images are randomly sampled and annotated from the generated images of nuclei of ROI1. The generated annotation file are shuffled and processed to generate a new label for rest of the nuclei to improve the performance of the network. 10% of the data is used as a testing set for cross validation purposes.

### 4.2 Convolutional Neural Networks

Convolutional neural networks (CNNs) are feed-forward, deep learning architectures made up of several connected convolutional layers [51, 52, 53]. CNNs approximate some underlying unknown function that maps input data to some target domain. In this case of this study, the mapping to be approximated is that of raw input image data to the classification of each pixel to a particular label (chromosome, nucleolus, etc.); i.e. supervised semantic segmentation.

As information passes through the network, each convolutional layer convolves the preceding layer's output with an increasing number of two-dimensional convolutional filters, resulting in an intermediate feature map that is passed along as input to the next layer or operation. The additional operations used between adjacent convolutional layers typically consist of nonlinear activation functions and normalization layers which help expedite the learning process, and max pooling to introduce translation invariance [54] and reduce computational costs via spatial coordinate downsampling.

Imperative to the CNN learning process are the convolutional filters. Each filter, typically of size  $3 \times 3$  or  $5 \times 5$ , consists of weights to be learned during network training and acts as a smaller receptive field of view that houses some learned feature from the overall image set which may then be re-used and applied to more-complex image reconstruction tasks in the later network layers. This allows for a more-global learning paradigm and deeper, more-parameter efficient neural network architectures than that of more-traditional fully-connected neural networks (FCNNs) [55, 56, 57] in which each individual pixel and connected node is assigned a learnable weight, and individual learned features remain localized to the single spatial coordinates in which they were found, not to be reused anywhere else in the network.

**U-Net** The main neural network model we implement for nucleolus segmentation is the U-Net [58], a deep convolutional network first used for pixel-by-pixel segmentation of biomedical images [58]. Inspired by convolutional autoencoders [59, 60], the U-Net, pictured in Fig. 3, is a symmetric encoder-decoder system made up of two distinct halves: the beginning contractive encoder-half on the left of Fig. 3 aims to capture contextual information and detect important image features with an increase in the number of convolutional channels and corresponding filters, while the expansive decoder-half on the right projects the learned features back into the higher resolution image space to reconstruct the input and predict a pixel-by-pixel semantic segmentation. Resulting from the encoder's contractive operations and partitioning the two halves is a compressed, lower-dimensional “bottleneck” which forces the network to learn a compression of the overall data and learn those features most imperative to the decoder reconstruction predictions.

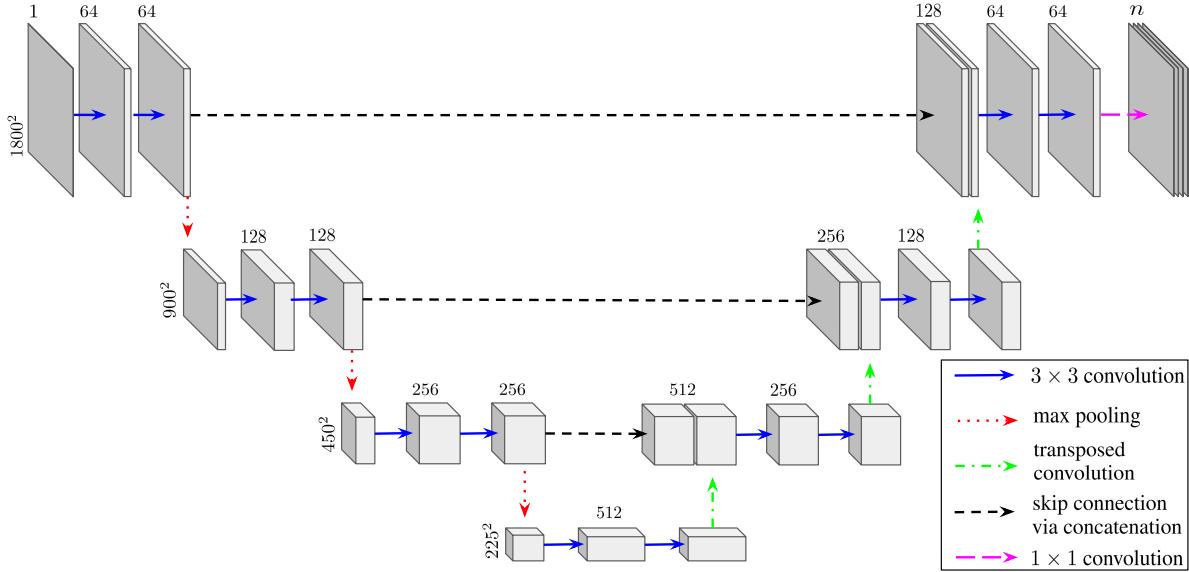


Figure 3: Schematic of a four-layer deep U-Net showing individual operations and intermediate feature map dimensions. In the left encoder-half, max pooling operations (red dotted arrows) halve spatial dimensions and convolutional operations (blue arrows) double the number of channels and filters at each subsequent layer. In the right decoder-half, convolutions decrease the number of channels and transposed convolutions (green dashed-dotted arrows) upsample the spatial dimensions. Skip connections (horizontal dashed arrows) join the two network halves. Lastly a single convolution with filter size  $1 \times 1$  (purple long dashed arrows) reduces the network output to any desired  $n$ -number of channels, which in this case is four (one for background and three for the nucleolus, chromosomes, and synaptonemal complex).

The U-Net remains popular in a number of current segmentation applications due to its robustness, simplicity, and ability to more readily propagate contextual information through the entirety of the network [61, 62]. This is accomplished through three means: a) an increase in the number of convolutional channels over traditional FCNNs, largely due to depth that U-Nets achieve, b) successive max-pooling of the data between network layers local features more-easily correlated with behavior and context at differing length scales [63], and c) channel-wise concatenations of encoder

feature map outputs to the decoder layers. These long-reaching concatenations, known as skip connections, decouple the encoder and decoder halves, allowing for an aggregation of multi-scale feature representation at different network stages [64, 65, 63] and helping alleviate the vanishing gradient problem which plagues deeper networks [66].

**Mixed-Scale Dense Networks** While U-net architectures remain popular, common implementations often require over several million trainable parameters which can lead to overfitting problems and harm network robustness, especially in applications where the amount of training data is low [56, 67]. In response, the MSDNet [68, 69] architecture, depicted in Fig. 4, was developed as a deep learning framework containing fewer trainable parameters (typically two to three orders of magnitude *fewer*) than U-Nets. This is accomplished by densely connecting *all* network layers to encourage maximum reusability of image features and by replacing the typical scaling operations found in encoder-decoder networks with dilated convolutions [70] in order to probe images at different length scales. By assigning a specific dilation to each MSDNet layer, the network can learn which dilation combinations are most effective. As a result, the number of network layers and the maximum integer dilation in which to cycle through are the most significant hyperparameters in which to toggle, drastically simplifying network design. Additionally, the dense connections among intermediate feature maps creates skip connections of *all* possible lengths. Lost spatial information is more readily recovered with the inclusion of these dense skip connections, which furthermore helps alleviate the vanishing gradient problem that plagues deep networks [66].

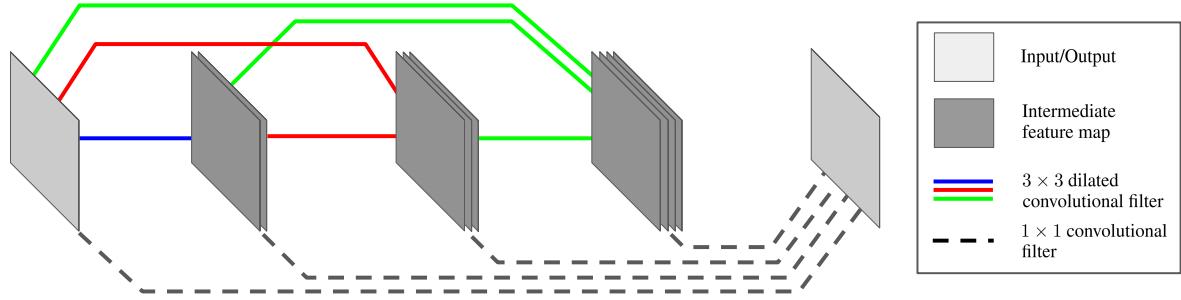


Figure 4: Schematic of a three-layer mixed-scale dense network (MSDNet). Blue, green, and red lines above represent  $3 \times 3$  dilated convolutions between each possible pair of input and intermediate layers. Different dilations are assigned to each color. Black dotted lines below represent the  $1 \times 1$  convolutional operator between the output and all other layers, amounting to a linear combination with learnable weights of all input and intermediate feature maps.

### 4.3 Training Parameters

During the model training phase, we use the multi-class cross entropy loss metric to measure how well the models classify each pixel to its respective class. The popular ADAM optimizer [71] was chosen to minimize the loss and update the neural network weights accordingly. As for the network learning rates, all neural networks were trained for 200 epochs with an initial rate of  $10^{-1}$  that was dropped by a factor of ten midway through training. For each trained model, a subset of 10% of training data was set aside as a validation set for cross-correlation purposes and to monitor model overfitting. The network weight set correlating to the epoch with the lowest validation set loss was chosen. Lastly, each model was trained on single Nvidia RTX 3090 GPU with 24 GB memory capacity and 936 GB/second bandwidth using a 10 core/20 thread Intel i9-10900X CPU.

### 4.4 Evaluation Metrics

To gauge model segmentation predictions in both the training and validation data sets, we use the F1 score, a popular measure of classifier performance [72]. It is defined as the harmonic mean of the model prediction's precision and recall, given by:

$$F1 = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}. \quad (1)$$

To further elaborate on this metric, it is useful to focus on the model predictions within each of the individual classes; we denote  $TP_i$  (true positive) as the number of pixels within a class  $i$  correctly identified by the model,  $FP_i$  (false positive) as the number of pixels incorrectly predicted to belong in class  $i$ , and  $FN_i$  (false negative) as the number of

pixels belonging to class  $i$  that are misclassified by the model. The confusion matrix diagrams these entities in Fig. 5. The model precision and recall within a single class  $i$  is then given by:

$$\text{precision}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i}, \quad \text{recall}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i}. \quad (2)$$

Precision and recall are often at odds with each other, as increasing recall (the ratio of how many instances within a particular ground truth class were correctly predicted) often reduces precision (the accuracy among all model predictions made of a single class), and vice-versa. For example, an overzealous classifier may over-predict, correctly identifying most instances of certain class but erroneously producing many more false positives, leading to suitable recall but poor precision. To alleviate this, the F1 metric offers a suitable balance between the two.

		Model Prediction	
		Positive	Negative
Ground Truth	Positive	TP	FN
	Negative	FP	TN

Figure 5: Confusion matrix of model predictions vs. the actual ground truth labels. Constituents of precision are highlighted vertically in purple, while constituents of recall are highlighted horizontally in red.

To calculate the F1 score for the entire model, individual F1 scores are calculated for each individual class from their respective precision and recall metrics in Eq. 2. The full model F1 score, our target evaluation metric in Eq. 1, then results from averaging each individual class score. To adjust for class imbalance, we compute the micro F1 score which aggregates the class scores by weighing each one differently according to relative size; i.e. the ratio of pixels belonging in each class to the total number of pixels in the dataset.

#### 4.5 Software Availability

Niraj: add a few sentences linking the reader to the github repo. Additionally, the U-Net and MSDNet models used here have been enhanced by allowing the specification of network architecture-defining hyperparameters, such as the number of network layers, initial number of channels, convolutional channel growth rate, and custom sets of MSDNet dilations. This level of user-defined custom implementation was accomplished through the Python-based *pyMSDtorch* deep learning software library (<https://pymsdtorch.readthedocs.io/>), which allows one to easily tune network hyperparameters to optimize performance.

## 5 Overall Workflow

The ROI1, ROI2, and ROI3 nuclei tiff stacks were divided into a single nuclei jpeg file. These images were pre-processed and annotated according to Sect. 4.1. Only nuclei images from ROI1 are used for training the various networks. This was due to the contrast between the three stacks being remarkably similar. Included in the first set of training data are four total annotations: one for the nucleolus, another for the chromosomes, and two background annotations, namely the interior nucleus background and exterior background located outside of the nucleus. The inclusion of two background classes greatly improved nucleolus and chromosome segmentation, as the exterior background class had a vastly different contrast and could be ignored in the gradient calculations and model parameter update steps. For the second segmentation model, the synaptonemal complex class was predicted, with prior synaptonemal complex annotations were added to the aforementioned four annotations. For the third and final network, the nucleolus and nuclear membrane were the only background elements in the third batch of training data. Lower accuracy is obtained by training a single network with training data that includes more classes membrane or synaptonemal complex. Choosing many networks for various classes therefore improves segmentation accuracy and yields better results. Lastly, once network training was complete and new segmentation predictions were inferred from the trained network, we performed

a post-processing step using skimage [73], a Python-based open-source image library, to filter out small, superfluous objects from each individual class. More specifically, small objects of volume 1000, 2000, and 3000 pixels were filtered out for the synaptonemal complex, chromosome, and nucleolus classes, respectively.

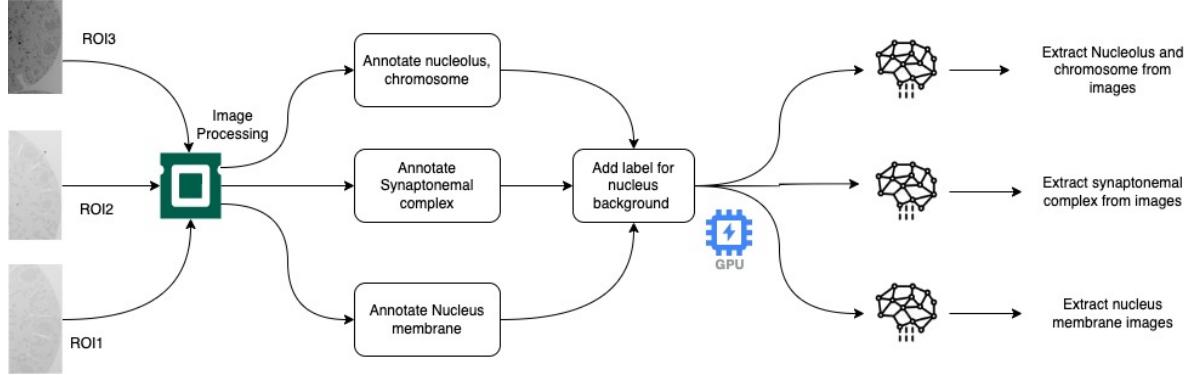


Figure 6: Complete end to end workflow for the segmentation of FIBSEM data.

Model	0	1	2	3	4
Depth	4	4	4	5	5
Base channels	32	64	64	32	64
Growth rate	2	2	1.5	2	2
Parameter count (millions)	2.143	8.560	2.57	8.63	34.512
Batch size	1	1	1	1	1
Training loss ( $10^{-2}$ )	2.86	2.94	2.97	2.73	2.78
Validation loss ( $10^{-2}$ )	3.19	3.3	3.42	3.40	3.45
Micro F1 score	.9015	.9023	.8958	.8984	.9041

Model	0	1	2	3	4
Depth	4	5	4	4	5
Base channels	32	32	64	64	64
Growth rate	2	2	1.5	2	2
Parameter count (millions)	2.143	8.630	2.57	8.56	34.512
Batch size	1	1	1	1	1
Training loss ( $10^{-2}$ )	3.617	3.82	3.89	3.55	3.41
Validation loss ( $10^{-2}$ )	3.988	3.88	4.105	3.97	3.87
Micro F1 score	.7657	.7808	.7446	.7687	0.7801

Table 1: The first table represents the results of UNET trained on data with nucleolus chromosome only whereas The second table represents the results for the UNET trained on the nucleolus chromosome along with synaptonemal complex. Best performing network is highlighted in gray.

## 6 Results

All networks were trained on an NVIDIA Ge Force RTX 3090 with 24GB of RAM. The number of network parameters that could be trained was restricted to around 40 million with an image quality of 1700x1700 and a batch size of 1. All of the networks are primarily trained on ROI1 images using random samples from various nuclei in ROI1. This is done to guarantee that all of the image contrast is used to train the network. Figure 7 depicts the segmentation results for ROI1 from the various trained networks. The first graphic depicts a 3D depiction of the first network's chromosome segmentation result. The second image is a segmented synaptonemal complex from the second network. The third image shows the nuclear membrane's 3D segmentation output from the third network. When compared to a network trained on all labels at once, training multiple networks for various labels yields better results. Table 1 shows the F1 scores for various UNETs trained on FIBSEM data with nucleolus, chromosome, and background labels, whereas Table 2 shows different scores for UNETs trained with an additional label of synaptonemal complex in addition to the current labels.

According to the tables, there is a considerable difference in accuracy with the addition of the label of synaptonemal complex. The variation in depth or base channels has no effect on training or validation losses in the first network. With a training F1 score of 0.9041, UNET with 64 base channels and a depth of 5 produces the best results. In contrast, variation in base channels and depth has a considerable impact on segmentation results and F1 scores in the second network. With a score of 0.7808, the network with 32 base channels and depth 5 performs better than the other networks. We extract chromosome and nucleolus from the first network and only the synaptonemal complex from the second network since the score for the first network is 0.9041. The third network, whose scores are identical to those of the network trained with synaptonemal complex, is used to extract the nuclear membrane. Figure 9 displays the nuclear membrane of ROI1 after 3D segmentation. The individual chromosomal divides that result from segmentation are shown in Figure 10. Individual connected chromosomes are labeled using 3D connected components [74]. By setting connectivity to 26 and delta to 10, the segmented chromosome's 12 biggest components can be identified using the cc3d library. Any nearby voxel value that is less than 10 is treated as the same component when delta is set.

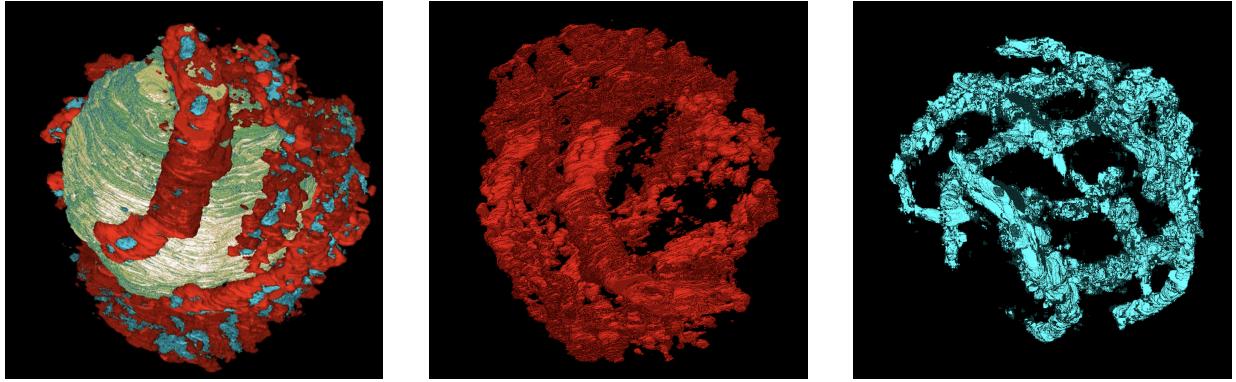


Figure 7: ROI1 stack segmentation results. The first image displays the 3d orientation of nucleolus, chromosome and synaptonemal complex in a nucleus. The synaptonemal complex is seen in the Third image, whereas the second image shows a three-dimensional representation of a chromosome.

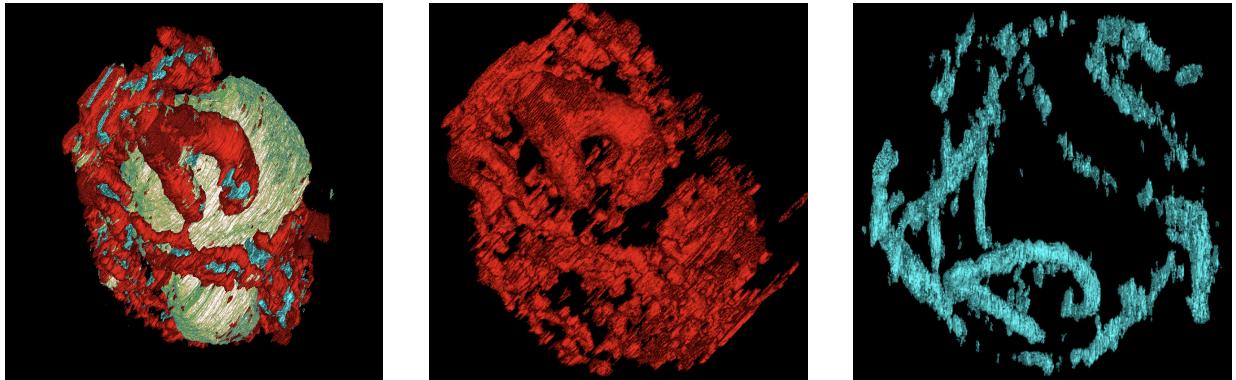


Figure 8: ROI2 stack segmentation results. The first image displays the 3d orientation of nucleolus, chromosome and synaptonemal complex in a nucleus. The synaptonemal complex is seen in the Third image, whereas the second image shows a three-dimensional representation of a chromosome.

## 7 Discussion and Future Work

This method compares favorably to other deep learning segmentation results for FIB-SEM data...

For future work, we envision deep learning models incorporating full 3D volumes of images. Though memory intensive, 3D deep learning methods may be able to better contextualize local information given the extra neighboring information in higher dimensionality. A challenge here remains, namely the difficulty in curating 3D masks and labels for a large and representative training set of images. However, this difficulty may be alleviated with the use of mixed-scale dense networks (MSDNets, detailed in Sect. 4.2). Though MSDNets performed sub-optimally compared to U-Nets in this current study, their densely-connected architecture and maximum reusability of data was specifically designed to

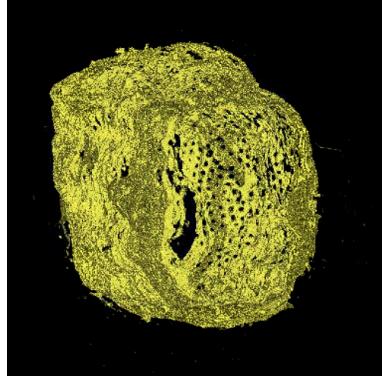


Figure 9: Segmentation results of nuclear membrane from ROI1.

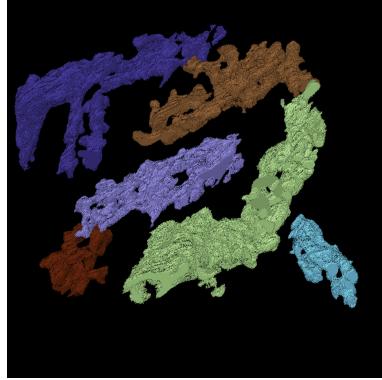


Figure 10: Extraction of individual chromosome from ROI1 using the 3d connected components.

perform better on smaller training data sets and sparse labeling [68, 69], and their ability to accommodate 3D volumes of images with little network architecture is advantageous. Alternatively, patch-based (or tile-based) data augmentation schemes in which smaller, overlapping subsets of images are drawn from the original data [75, 76] may be generalized to 3D volumes of data, allowing 3D U-Nets and MSDNets to learn from significantly smaller sets of labels and training data, evidenced by similar overlap averaging techniques [77].

## 8 Funding

Further support originates from the National Institute of General Medical Sciences of the National Institutes of Health (NIH) under Award 5R21GM129649-02 and from the Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory under U.S. Department of Energy Contract No. DE-AC02-05CH11231. **For Peter: please double check these.**

## 9 Acknowledgements

Some people or groups to add?

## 10 Author Contributions

VK and PHZ supervised the project; NG, EJR, SP, CSX, HFH, and VK wrote the manuscript (original draft); VK, PHZ, NG, and EJR reviewed and edited the manuscript; CSX developed the enhanced FIB-SEM platform for the high resolution imaging; CSX, SP, and HFH conducted the FIB-SEM imaging of *C. elegans*; CSX, SP, and HFH prepared *C. elegans* samples; NG and EJR performed data analysis and machine learning architecture design; EJR and PHZ designed pyMSDtorch machine learning software suite; NG and VK uploaded data and provided workflow to GitHub repository; VK proposed the biological questions; PHZ proposed the machine learning solutions.

## References

- [1] Yumei Wu, Christina Whiteus, C Shan Xu, Kenneth J Hayworth, Richard J Weinberg, Harald F Hess, and Pietro De Camilli. Contacts between the endoplasmic reticulum and other membranes in neurons. *Proceedings of the National Academy of Sciences*, 114(24):E4859–E4867, 2017.
- [2] KJ Hayworth, N Kasthuri, R Schalek, and JW Lichtman. Automating the collection of ultrathin serial sections for large volume tem reconstructions. *Microscopy and Microanalysis*, 12(S02):86–87, 2006.
- [3] AA Wanner, MA Kirschmann, and C Genoud. Challenges of microtome-based serial block-face scanning electron microscopy in neuroscience. *Journal of microscopy*, 259(2):137–142, 2015.
- [4] Winfried Denk, Heinz Horstmann, and Kristen M Harris. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS biology*, 2(11):e329, 2004.
- [5] RK Nalla, AE Porter, C Daraio, AM Minor, V Radmilovic, EA Stach, AP Tomsia, and RO Ritchie. Ultrastructural examination of dentin using focused ion-beam cross-sectioning and transmission electron microscopy. *Micron*, 36(7-8):672–680, 2005.
- [6] Lucille A Giannuzzi and Frederick A Stevie. A review of focused ion beam milling techniques for tem specimen preparation. *Micron*, 30(3):197–204, 1999.
- [7] Lorenz Holzer, Fedir Indutnyi, PH Gasser, Beat Münch, and Markus Wegmann. Three-dimensional analysis of porous batio3 ceramics using fib nanotomography. *Journal of microscopy*, 216(1):84–95, 2004.
- [8] Shaina Kelly, Hesham El-Sobky, Carlos Torres-Verdín, and Matthew T Balhoff. Assessing the utility of fib-sem images for shale digital rock physics. *Advances in water resources*, 95:302–316, 2016.
- [9] Hao Wu, Yanbin Yao, Yingfang Zhou, and Feng Qiu. Analyses of representative elementary volume for coal using x-ray  $\mu$ -ct and fib-sem and its application in permeability predication model. *Fuel*, 254:115563, 2019.
- [10] Jeff T Gostick, Zohaib A Khan, Thomas G Tranter, Matthew DR Kok, Mehrez Agnaou, Mohammadamin Sadeghi, and Rhodri Jervis. Porespy: A python toolkit for quantitative analysis of porous media images. *Journal of Open Source Software*, 4(37):1296, 2019.
- [11] Huihuang Fang, Shuxun Sang, Shiqi Liu, and Yi Du. Methodology of three-dimensional visualization and quantitative characterization of nanopores in coal by using fib-sem and its application with anthracite in qinshui basin. *Journal of Petroleum Science and Engineering*, 182:106285, 2019.
- [12] Mohamed Garum, Paul WJ Glover, Piroska Lorinczi, Rik Drummond-Brydson, and Ali Hassanpour. Micro-and nano-scale pore structure in gas shale using  $x\mu$ -ct and fib-sem techniques. *Energy & Fuels*, 34(10):12340–12353, 2020.
- [13] Martin Čalkovský, Erich Müller, Matthias Meffert, Nadejda Firman, Frederik Mayer, Martin Wegener, and Dagmar Gerthsen. Comparison of segmentation algorithms for fib-sem tomography of porous polymers: Importance of image contrast for machine learning segmentation. *Materials Characterization*, 171:110806, 2021.
- [14] Magnus Röding, C Fager, A Olsson, C Von Corswant, E Olsson, and Niklas Loren. Three-dimensional reconstruction of porous polymer films from fib-sem nanotomography data using random forests. *Journal of Microscopy*, 281(1):76–86, 2021.
- [15] Cecilia Fager, Magnus Röding, Anna Olsson, Niklas Lorén, Christian von Corswant, Aila Särkkä, and Eva Olsson. Optimization of fib-sem tomography and reconstruction for soft, porous, and poorly conducting materials. *Microscopy and Microanalysis*, 26(4):837–845, 2020.
- [16] Damjana Drobne, Marziale Milani, Monica Ballerini, Alexis Zrimec, Maja Berden Zrimec, Francesco Tatti, and Kazimir Draslar. Focused ion beam for microscopy and in situ sample preparation: application on a crustacean digestive system. *Journal of Biomedical Optics*, 9(6):1238–1243, 2004.
- [17] Caroline Kizilyaprak, York-Dieter Stierhof, and Bruno M Humbel. Volume microscopy in biology: Fib-sem tomography. *Tissue and Cell*, 57:123–128, 2019.
- [18] David P Hoffman, Gleb Shtengel, C Shan Xu, Kirby R Campbell, Melanie Freeman, Lei Wang, Daniel E Milkie, H Amalia Pasolli, Nirmala Iyer, John A Bogovic, et al. Correlative three-dimensional super-resolution and block-face electron microscopy of whole vitreously frozen cells. *Science*, 367(6475):eaaz5357, 2020.
- [19] Andreas Müller, Deborah Schmidt, C Shan Xu, Song Pang, Joyson Verner D’Costa, Susanne Kretschmar, Carla Münster, Thomas Kurth, Florian Jug, Martin Weigert, et al. 3d fib-sem reconstruction of microtubule–organelle interaction in whole primary mouse  $\beta$  cells. *Journal of Cell Biology*, 220(2), 2021.

- [20] Aubrey V Weigel, Chi-Lun Chang, Gleb Shtengel, C Shan Xu, David P Hoffman, Melanie Freeman, Nirmala Iyer, Jesse Aaron, Satya Khuon, John Bogovic, et al. Er-to-golgi protein delivery through an interwoven, tubular network extending from er. *Cell*, 184(9):2412–2429, 2021.
- [21] Damjana Drobne, M Milani, A Zrimec, V Lešer, and M Berden Zrimec. Electron and ion imaging of gland cells using the fib/sem system. *Journal of microscopy*, 219(1):29–35, 2005.
- [22] Netta Vidavsky, Anat Akiva, Ifat Kaplan-Ashiri, Katya Rechav, Lia Addadi, Steve Weiner, and Andreas Schertel. Cryo-fib-sem serial milling and block face imaging: Large volume structural analysis of biological tissues preserved close to their native state. *Journal of Structural Biology*, 196(3):487–495, 2016.
- [23] V González-Ruiz, JP García-Ortiz, MR Fernández-Fernández, and JJ Fernández. Optical flow driven interpolation for isotropic fib-sem reconstructions. *Computer Methods and Programs in Biomedicine*, 221:106856, 2022.
- [24] Tim Dahmen, Michael Engstler, Christoph Pauly, Patrick Trampert, Niels De Jonge, Frank Mücklich, and Philipp Slusallek. Feature adaptive sampling for scanning electron microscopy. *Scientific reports*, 6(1):1–11, 2016.
- [25] Larissa Heinrich, Davis Bennett, David Ackerman, Woohyun Park, John Bogovic, Nils Eckstein, Alyson Petruncio, Jody Clements, Song Pang, C Shan Xu, et al. Whole-cell organelle segmentation in volume electron microscopy. *Nature*, 599(7883):141–146, 2021.
- [26] C Shan Xu, Song Pang, Gleb Shtengel, Andreas Müller, Alex T Ritter, Huxley K Hoffman, Shin-ya Takemura, Zhiyuan Lu, H Amalia Pasolli, Nirmala Iyer, et al. An open-access volume electron microscopy atlas of whole cells and tissues. *Nature*, 599(7883):147–151, 2021.
- [27] Alex J Perez, Mojtaba Seyedhosseini, Thomas J Deerinck, Eric A Bushong, Satchidananda Panda, Tolga Tasdizen, and Mark H Ellisman. A workflow for the automatic segmentation of organelles in electron microscopy image stacks. *Frontiers in neuroanatomy*, 8:126, 2014.
- [28] Matthew Andrew. A quantified study of segmentation techniques on synthetic geological xrm and fib-sem images. *Computational Geosciences*, 22(6):1503–1512, 2018.
- [29] Sydney Brenner. The genetics of caenorhabditis elegans. *Genetics*, 77(1):71–94, 1974.
- [30] Jonathan A Hodgkin and Sydney Brenner. Mutations causing transformation of sexual phenotype in the nematode caenorhabditis elegans. *Genetics*, 86(2):275–287, 1977.
- [31] Alan Coulson, John Sulston, Sydney Brenner, and Jonathan Karn. Toward a physical map of the genome of the nematode caenorhabditis elegans. *Proceedings of the National Academy of Sciences*, 83(20):7821–7825, 1986.
- [32] C. elegans Sequencing Consortium\*. Genome sequence of the nematode c. elegans: a platform for investigating biology. *Science*, 282(5396):2012–2018, 1998.
- [33] Martin Chalfie, John E Sulston, John G White, Eileen Southgate, J Nicol Thomson, and Sydney Brenner. The neural circuit for touch sensitivity in caenorhabditis elegans. *Journal of Neuroscience*, 5(4):956–964, 1985.
- [34] Diya Banerjee, Alvin Kwok, Shin-Yi Lin, and Frank J Slack. Developmental timing in c. elegans is regulated by kin-20 and tim-1, homologs of core circadian clock genes. *Developmental cell*, 8(2):287–295, 2005.
- [35] Harris S Kaplan, Oriana Salazar Thula, Niklas Khoss, and Manuel Zimmer. Nested neuronal dynamics orchestrate a behavioral hierarchy across timescales. *Neuron*, 105(3):562–576, 2020.
- [36] John G White, Eileen Southgate, J Nichol Thomson, Sydney Brenner, et al. The structure of the nervous system of the nematode caenorhabditis elegans. *Philos Trans R Soc Lond B Biol Sci*, 314(1165):1–340, 1986.
- [37] Emma K Towlson, Petra E Vértes, Sebastian E Ahnert, William R Schafer, and Edward T Bullmore. The rich club of the c. elegans neuronal connectome. *Journal of Neuroscience*, 33(15):6380–6387, 2013.
- [38] Steven J Cook, Travis A Jarrell, Christopher A Brittin, Yi Wang, Adam E Bloniarz, Maksim A Yakovlev, Ken CQ Nguyen, Leo T-H Tang, Emily A Bayer, Janet S Duerr, et al. Whole-animal connectomes of both caenorhabditis elegans sexes. *Nature*, 571(7763):63–71, 2019.
- [39] C Shan Xu, Kenneth J Hayworth, Zhiyuan Lu, Patricia Grob, Ahmed M Hassan, José G García-Cerdán, Krishna K Niyogi, Eva Nogales, Richard J Weinberg, and Harald F Hess. Enhanced fib-sem systems for large-volume 3d imaging. *Elife*, 6:e25916, 2017.
- [40] C Shan Xu, Kenneth J Hayworth, and Harald F Hess. Enhanced fib-sem systems for large-volume 3d imaging, March 24 2020. US Patent 10,600,615.
- [41] David G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision*, volume 2, pages 1150–1157. Ieee, 1999.
- [42] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

- [43] Heng Yang, Guojian Cheng, and Hongchu Chen. High efficient local feature matching. In *2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pages 2552–2556. IEEE, 2018.
- [44] Pyushi Singhal, Akhilesh Verma, and Akash Garg. A study in finding effectiveness of gaussian blur filter over bilateral filter in natural scenes for graph based image segmentation. In *2017 4th international conference on advanced computing and communication systems (ICACCS)*, pages 1–6. IEEE, 2017.
- [45] Jean Serra. Image analysis and mathematical morphol-ogy. (*No Title*), 1982.
- [46] Jean-Francois Rivest, Pierre Soille, and Serge Beucher. Morphological gradients. *Journal of Electronic Imaging*, 2(4):326–336, 1993.
- [47] James Lee, R Haralick, and Linda Shapiro. Morphologic edge detection. *IEEE Journal on Robotics and Automation*, 3(2):142–156, 1987.
- [48] Jufriadiif Na’am, Johan Harlan, Rosda Syelly, and Agung Ramadhanu. Filter technique of medical image on multiple morphological gradient (mmg) method. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 17(3):1317–1323, 2019.
- [49] A Etemadi. Robust segmentation of edge data. In *1992 International Conference on Image Processing and its Applications*, pages 311–314. IET, 1992.
- [50] Cuneyt Akinlar and Cihan Topal. Edlines: A real-time line segment detector with a false detection control. *Pattern Recognition Letters*, 32(13):1633–1642, 2011.
- [51] Kunihiko Fukushima and Sei Miyake. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets*, pages 267–285. Springer, 1982.
- [52] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [53] Patrice Y Simard, David Steinkraus, John C Platt, et al. Best practices for convolutional neural networks applied to visual document analysis. In *Icdar*, volume 3. Edinburgh, 2003.
- [54] Dominik Scherer, Andreas Müller, and Sven Behnke. Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks*, pages 92–101. Springer, 2010.
- [55] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.
- [56] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.
- [57] Qi Xu, Ming Zhang, Zonghua Gu, and Gang Pan. Overfitting remedy by sparsifying regularization on fully-connected layers of cnns. *Neurocomputing*, 328:69–74, 2019.
- [58] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.
- [59] David E. Rumelhart and James L. McClelland. *Learning Internal Representations by Error Propagation*, pages 318–362. 1987.
- [60] David DeMers and Garrison Cottrell. Non-linear dimensionality reduction. *Advances in neural information processing systems*, 5, 1992.
- [61] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.
- [62] Narinder Singh Punn and Sonali Agarwal. Modality specific u-net variants for biomedical image segmentation: a survey. *Artificial Intelligence Review*, pages 1–45, 2022.
- [63] Kyoung Jin Noh, Sang Jun Park, and Soochahn Lee. Scale-space approximated convolutional neural networks for retinal vessel segmentation. *Computer methods and programs in biomedicine*, 178:237–246, 2019.
- [64] Pulkit Kumar, Pravin Nagar, Chetan Arora, and Anubha Gupta. U-segnet: fully convolutional neural network based automated brain tissue segmentation tool. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3503–3507. IEEE, 2018.

- [65] Michal Drozdzal, Eugene Vorontsov, Gabriel Chartrand, Samuel Kadoury, and Chris Pal. The importance of skip connections in biomedical image segmentation. In *Deep learning and data labeling for medical applications*, pages 179–187. Springer, 2016.
- [66] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR, 2015.
- [67] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.
- [68] Daniël M Pelt and James A Sethian. A mixed-scale dense convolutional neural network for image analysis. *Proceedings of the National Academy of Sciences*, 115(2):254–259, 2018.
- [69] Daniël M Pelt, Kees Joost Batenburg, and James A Sethian. Improving tomographic reconstruction from limited data using mixed-scale dense convolutional neural networks. *Journal of Imaging*, 4(11):128, 2018.
- [70] Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.
- [71] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [72] Nancy Chinchor and Beth M Sundheim. Muc-5 evaluation metrics. In *Fifth Message Understanding Conference (MUC-5): Proceedings of a Conference Held in Baltimore, Maryland, August 25-27, 1993*, 1993.
- [73] Stefan Van der Walt, Johannes L Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014.
- [74] William Silversmith. cc3d: Connected components on multilabel 3d images. *Nature Methods*, 2021.
- [75] Carlo Innamorati, Tobias Ritschel, Tim Weyrich, and Niloy J Mitra. Learning on the edge: Investigating boundary filters in cnns. *International Journal of Computer Vision*, pages 1–10, 2019.
- [76] Yuxin Cui, Guiying Zhang, Zhonghao Liu, Zheng Xiong, and Jianjun Hu. A deep learning algorithm for one-step contour aware nuclei segmentation of histopathology images. *Medical & biological engineering & computing*, 57(9):2027–2043, 2019.
- [77] Nicolas Pielawski and Carolina Wählby. Introducing hann windows for reducing edge-effects in patch-based image segmentation. *PloS one*, 15(3):e0229839, 2020.