Learning to Correct Sloppy Annotations in Electron Microscopy Volumes

Minghao Chen[1], Mukesh Bangalore Renuka[2], Lu Mi[1], Jeff Lichtman[2], Nir Shavit[1], Yaron Meirovitch[2]
MIT[1], Harvard University[2]

Abstract

Connectomics deals with the problem of reconstructing neural circuitry from electron microscopy images at the synaptic level. Automatically reconstructing circuits from these volumes requires high fidelity 3-D instance segmentation, which yet appears to be a daunting task for current computer vision algorithms. Hence, to date, most datasets are not reconstructed by fully-automated methods. Even after painstaking proofreading, these methods still produce numerous small errors. In this paper, we propose an approach to accelerate manual reconstructions by learning to correct imperfect manual annotations. To achieve this, we designed a novel solution for the canonical problem of marker-based 2-D instance segmentation, reporting a new state-of-the-art for region-growing algorithms demonstrated on challenging electron microscopy image stacks. We use our marker-based instance segmentation algorithm to learn to correct all "sloppy" object annotations by reducing and expanding all annotations. Our correction algorithm results in high quality morphological reconstruction (near ground truth quality), while significantly cutting annotation time (∼8x) for several examples in connectomics. We demonstrate the accuracy of our approach on public connectomics benchmarks and on a set of large-scale neuron reconstruction problems, including on a new octopus dataset that cannot be automatically segmented at scale by existing algorithms.

1. Introduction

Recent advances in electron microscopy (EM) allow scientists to image neurons at the synaptic level using nanoscale resolution. *Connectomics* is the field of neurobiology that aims to reconstruct the neural circuitry from these image-stacks, and ultimately from entire organisms [1, 75]. A cubic millimeter of brain tissue, the size of a grain of sand, contains about 100 thousand neurons, and is represented by a petabyte-scale image-stack of EM data. Data on such a scale cannot be reconstructed manually in reasonable time, and we must rely on high-fidelity and high-performance computer systems to effectively segment the objects (neurons) in the image volume [46]. This trend will be exacerbated as scientists aim for an exabyte scale complete mouse brain [1].

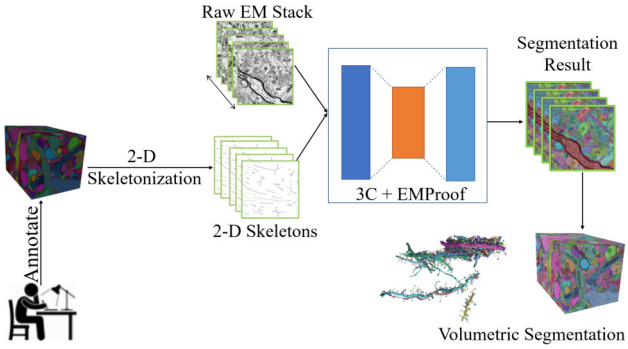

Figure 1. A high-level overview of our work. Swiftly made inaccurate annotations (imperfect instance segmentation) are corrected by a single computational primitive EMProof with end-to-end learning. The input to EMProof can be any partial instance segmentation such as skeletons or tiny seeds.

Connectomics thus poses one of the hardest known cases of automated 3-D instance segmentation. In other domains, 2-D instance segmentation tasks have received considerable attention, and satisfactory solutions exist for natural [13, 14, 19, 35, 47] and biomedical [20, 28, 67, 72, 73] images, and for video tracking applications [43, 71, 74]. However, there has not been a 3-D instance segmentation solution that can meet human reconstruction accuracy, and fully-automated methods are applicable only for datasets of exceptional quality (see progress by [4–6, 30, 39, 44, 51]).

The solution for small connectomics labs is thus to manually reconstruct small EM volumes. Occasionally, hybrid solutions are applied on larger volumes of high image quality: firstly perform automatic reconstruction based on machine learning algorithms, and then follow it with a manual human-expert led correction process. Unfortunately, because automated solutions make unacceptable amounts of errors on very thin objects even on high-quality datasets, the human corrector, even with state-of-the-art semi-automated proofreading tools such as [10, 16, 27, 33, 80], cannot effectively correct *all* errors and attain ground truth quality. Of-

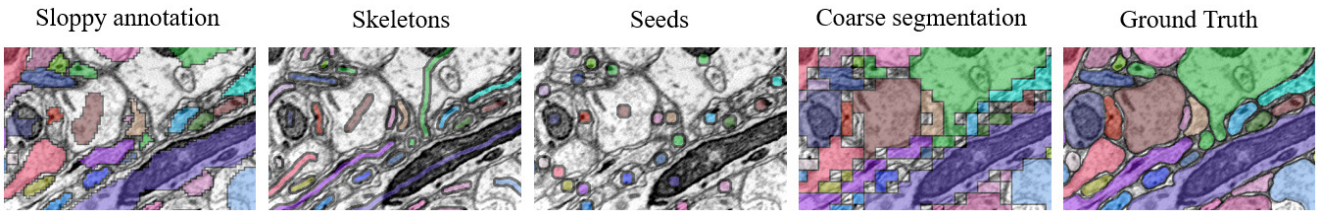

Figure 2. Inputs to EMProof: Sloppy annotations, Skeleton seeds, Round seeds and Coarse segmentation, compared with the ground truth.

ten, hundreds of undergraduate students [36] are required to painstakingly certify or correct every slice and every tile of a terabyte-scale 3-D reconstruction merely to get a wiring diagram. A recent study employed 87 annotators for about 4,000 human annotation hours [55] by proofreading thousands of thin objects of an automated reconstruction, which is a considerable improvement to the 20,000 annotation hours of an earlier study of a mammalian retina [36]. Although significant progress has been made to improve the proofreading speeds and accuracy [7, 25, 33, 49, 57, 61, 64], hybrid approaches so far have not been able to produce the level of accuracy obtained by pure manual segmentation. So far, no large scale volume has been densely segmented and proofread at the accuracy offered by purely manual segmentation [41]. Furthermore, since automated solutions are costly and require datasets of supreme segmentability [39, 40], numerous labs have focused on small datasets that can be manually reconstructed. Significantly cutting down human investment in manual annotation tasks is therefore of immense importance.

In this paper, we advocate an alternative approach for the cases where ground truth quality is required: have human annotators invest a minimal amount of time first, to quickly create "sloppy" annotations of EM images, and then use an automated machine learning technique to complete and correct these "sloppy" annotations into highly accurate reconstructions. A high level overview of our pipeline is described in Fig. 1. To achieve this ambitious goal, we set out to improve the state-of-the-art in seeded segmentation of electron microscopy images. We believe pushing forward this canonical problem could have numerous applications beyond connectomics.

For connectomics, we present a new technique for correcting "sloppy" annotations of EM images. "Sloppy" in this context refers to manual pixel paintings that are done hastily in as little time as possible as shown in Fig. 2, which tend not to respect neuron boundaries [54]. We design our solution based on the wisdom in connectomics that quick expert annotators tend to produce many geometrical errors while maintaining correct object topology. Our pipeline corrects these errors by leveraging and improving recent advances in label propagation across EM images [51]. We present an end-to-end learning mechanism, EMProof, which expands and corrects imperfect sloppy annotations. We investigate this strategy and show that it offers a segmentation with negligible morphological errors (negligible variation of information), significantly outperforming strong seed-growing techniques.

This suggests that a collaboration between a sloppy, but very fast human annotator, and an effective machine learning tool, can cut down human resources resulting in a much faster reconstruction compared to a purely manual one. Our pipeline also manages to attain accuracy akin to the ground truth. Since our pipeline relies on manual tracking of the objects, object splits and merge errors that are common to auto-segmentation pipelines, are not present in our solution. Our pipeline is also able to take human-annotated skeletons (as is often produced in connectomics; e.g., [3, 70, 76]) and map them directly to volumetric objects.

EMProof takes an imperfect 3-D annotation and aims to repaint it correctly using supervised learning methods. Imperfect annotations are first skeletonized in 2-D, followed by a novel fast segmentation algorithm that expands these skeleton objects to accurately attain the correct neuronal boundaries. We use skeletons rather than the annotations themselves to ensure that object's topology takes precedence. This reduces the impact of the incorrect painting over the neurons boundaries that are often made by the "sloppy" annotators. We also choose to go with the skeletons and not the annotations themselves to make the pipeline more invariant to the annotation and painting style of the annotators (which vary widely).

We demonstrate the scalability of our pipeline by accelerating the reconstruction time of a new large-scale *Octopus vulgaris* brain electron microscopy dataset [11], as well as on the widely studied mouse cortex dataset of [41], offering a speedup of up to 8x over the manual annotation with only marginal loss in quality. We demonstrate the accuracy of EMproof using the public benchmark CREMI [29] (fly brain) for 2-D instance segmentation. Although simpler, our solution provides a new state-of-the-art for this canonical problem. Most seeded image segmentation methods follow two steps [15, 17, 56, 65, 79]: first predicting the affinity graph of the image, and then diffusing the seeds from their initial locations into the foreground as defined by the affinity graph. We replace this practice by training a neu-

ral network, end-to-end, to predict labels directly, removing the dependency on affinity maps. Utilizing this simpler design while outperforming strong baselines on challenging EM tasks is likely to have direct impact on other biomedical problems.

We continue by demonstrating the ability of EMProof to achieve super-resolution of instance segmentation, in the context of connectomics, on two public connectomic datasets (AC3 and a larger-scale Somatosensory Cortex dataset [41]).

The main contributions of this paper are as follows:

- We improve the accuracy of the canonical problem of seeded segmentation and demonstrate it on challenging electron microscopy images, offering a strong alternative to watershed-like algorithms, while not requiring intermediate affinity maps.
- We propose a new way to accelerate high quality manual reconstruction of neurons from electron microscopy volumes (attaining ground truth quality on several data sets).
- We define a new class of problems dealing with the correction and proofreading of human annotations.

The rest of this paper is organized as follows. We review related work in Section 2 and describe the methodological issues in Section 3. The details of our experiments are provided in Section 4.

2. Related Work

Traditional approaches for seeded segmentation involved three steps. The first would be the generation of affinity graphs. These affinity graphs would then be used to define seeds which would then be propagated to obtain a seeded segmentation. Before the advent of deep learning, affinity graphs were obtained by computing domain specific feature maps such as color saliency [26, 60, 66] and random forest based constructions [24, 42, 84]. The pioneering work of Ciresan et al. [21], which like ours focused on connectomics, demonstrated that deep learning can produce superior border predictions [62]. Although, it still builds on a traditional design after this step by employing non-trainable algorithms such as diffusion on affinity graphs to expand seeds [44, 58]. Non-trainable algorithms, while successful in some cases [6, 12], cannot be seamlessly adapted for other use cases.

Several recent supervised learning studies show that learning to diffuse seeds in space is possible, for example, by learned watersheds [79] and learned random walkers [17] and their variants [77, 78]. These have performed exceedingly well; however, in similar to the model-based optimization studies, they require a preliminary border map or affinity graph on which learning to diffuse takes place.

The dependency of border maps is tackled by Flood Filling Networks (FFNs) by Januszewski et al. [39]. FFNs learn to expand object seeds to their appropriate territories in one step without the aid of border maps. The primary drawbacks of the algorithm are related to scalability as it operates on individual masks, thus requiring to iterate over thousands of objects, sometimes lying within a single image tile. Another drawback is that FFNs use complicated recursive learning and non-trivial heuristics. Cross Classification Clustering (3C) [51] improved the scalability of FFNs by offering an end-to-end approach to jointly classify multiple object instances, thereby improving scalability. However, 3C suffers from a dependency on high-quality border maps to define the initial seeds and propagate them across image sections. Our proposed pipeline adopts 3C as a computational primitive, whilst eliminating the dependency on border maps by using an attention network which jointly predicts both the borders and the segmentation results. Furthermore, we redesign the 3C primitive to operate with small seeds, improving segmentation resiliency and improving the accuracy of current seeded segmentation algorithms (see below).

Seeded Segmentation Seeded segmentation [17, 45, 79, 81] is a special case of instance segmentation where the objects labels are already associated with predefined sets of pixels (the seeds). Seeded segmentation is an important problem both for general instance segmentation tasks [22, 31] and for interactive segmentation [18, 52, 59]. Grady et al. [32] proposed the random walker algorithm which uses linear diffusion to solve the seeded segmentation problem. Zhu et al. associated Gaussian Markov Random Fields (GMRFs) with the problem of seeded segmentation [83]. Samuel et al. managed to learn context dependent potentials for GMRFs [63]. Jancsary et al. pioneered the use of flexible nonparmetric classifiers to determine the potentials in a Conditional GMRF [38]. Bertasius et al. attempted learning an end-to-end network for GMRF [9]. Cousty et al. proposed watershed method based on min-max criterion [23]. Although these methods have been used quite successfully, they are limited by the quality of the initially computed affinity graph. Our end-to-end trained pipeline eliminates the dependency on affinity graphs in seeded segmentation, and at the same time improving the accuracy of the expanded objects, which we believe could be useful to problems outside connectomics.

Attention Our initial experiments indicated that learning to expand very thin objects is hard to achieve, perhaps since reliably propagating information from feature masks over a long distance requires considerable network structure. We propose to tackle this difficulty by including an attention mechanism which has shown success in various domains [37, 69, 82]. Vaswani et al. used attention mechanisms to produce top scores in machine translation [69]. Later, Hu et al. proposed attention mecha-

nisms using Squeeze-and-Excitation Networks which use channel-wise relationships to enhance the representational power of the network [37]. In EMProof, we incorporate a self-attention block to capture the long-range relationship between the seeds and the pixels.

3. Methodology

In this section, we describe our automated correction of human-made neuronal annotations. We begin with a top-down description of our algorithm and continue with a description of our approach toward the end-to-end learning of seeded segmentation. We then describe our architecture and the self-attention mechanism used in our experiments while providing our design insights.

3.1. Top-down Description

Our general approach is described in Fig. 1. The pipeline takes as input a "sloppy" painting that is done by a quick annotator who paints hastily on each image. Our EMProof method learns how to map such an error-prone weak annotation into one that has correct morphology and crisp boundaries. Applying such transformation one object at a time is wasteful and will require to compute the neural network model a linear number of times (in the number of object instances). We use the idea of Cross Classification Clustering (3C) [51] as shown in Fig. 4 to perform this classification by computing the neural network *logarithmic number of times*. First, to alleviate the dependency on the style of annotation and keep it robust, we compute skeletal seeds for each of the instance objects in the image. We then colorize the labels of the skeletons using l-length strings using k different colors, where $l = \log(n)$ and n is the number of distinct skeleton labels in the input image. This yields l different skeleton images, where each image now only has k colors. Each pair of skeleton and EM images is then fed into the network. Each computation of the model will produce an image with k expanded skeletons. The 3C mechanism guarantees that the $l = \log(n)$ model predictions can be concatenated pixel-wise and uniquely remapped to the labels appearing in the input image (allowing arbitrary number of instances). Hence, all expanded skeleton images form together the final result (instance segmentation).To accomplish the above we redesigned 3C entirely so as to work with thin and/or round seeds (sparse inputs) rather than densely segmented images as in [51], and unlike other works to operate directly on microscopy images without affinity maps.

3.2. Architecture

The general network architecture, as given in Fig. 3, is a fully convolutional neural network [48] borrowing design principles from the widely used semantic segmentation network U-Net [62]. Our experiments supported replacing the encoder part of U-Net with the SE-ResNext-50 blocks [37]. In addition, the deepest representational features are transformed by self-attention block, where it encourages the feature tensor $X \in \mathbb{R}^{C \times W \times H}$ to focus closely on the input seeds and the object boundaries. After that, we split the network into two branches, segmentation and boundary prediction, to force the network to learn the spatial integrity of the instances. Each of the branches has the same architecture. The decoders take as input the aggregated information $O \in \mathbb{R}^{(C+C') \times W \times H}$ from a self-attention block. To capture multi-scale information, decoders are concatenated with an encoder using the same strategy of U-Net, and the upsampled outputs of the decoders are aggregated into a hypercolumn. Last, a fully convolutional layer is applied to make the prediction for each branch; one branch predicts object boundaries while the other predicts k-class instance segmentation. We present below results from ablation studies corroborating the above design choices.

3.3. Self Attention

The self attention mechanism has been shown to be successful in various machine learning tasks, such as machine translation, image analysis, video classification and question answering [37,69,82]. This is due to the fact that self attention helps the network to learn long-range dependencies and non-local features. Generally, an attention function is a mapping between the query, key-value pairs and outputs, which together inform the network where to "attend" when facing different queries. In EMProof, We modified the successful U-Net architecture using a self-attention module as shown in Fig. 3, in an attempt to force the network to focus on the input seeds and their topographic association to the far away boundaries.

4. Experiments

We evaluate the efficacy of EMProof in multiple ways. In the first experiment, we tested the usefulness of EMProof in correcting sloppy annotations made by experts on large-scale datasets. We present both the running time of EMProof and also an empirical study of a real annotator. Altogether, we present an 8-fold improvement in the nearly error-free reconstruction time, compared to the workload of experts.

While EMProof is used to perform 3-D instance corrections, it solves the problem of seeded segmentation along the way in 2-D. We aim to benchmark the performance of this core aspect of EMProof in order to demonstrate the accuracy and generality of the pipeline, compared to alternatives, and to aid other biomedical problems with a new effective technique for seeded segmentation. We test the expansion capabilities of EMProof both on small rounded seeds and on skeleton-like objects, which are the inputs to EMProof. We use the challenging CREMI benchmark and datasets for connectomics [2] to compare our solution to

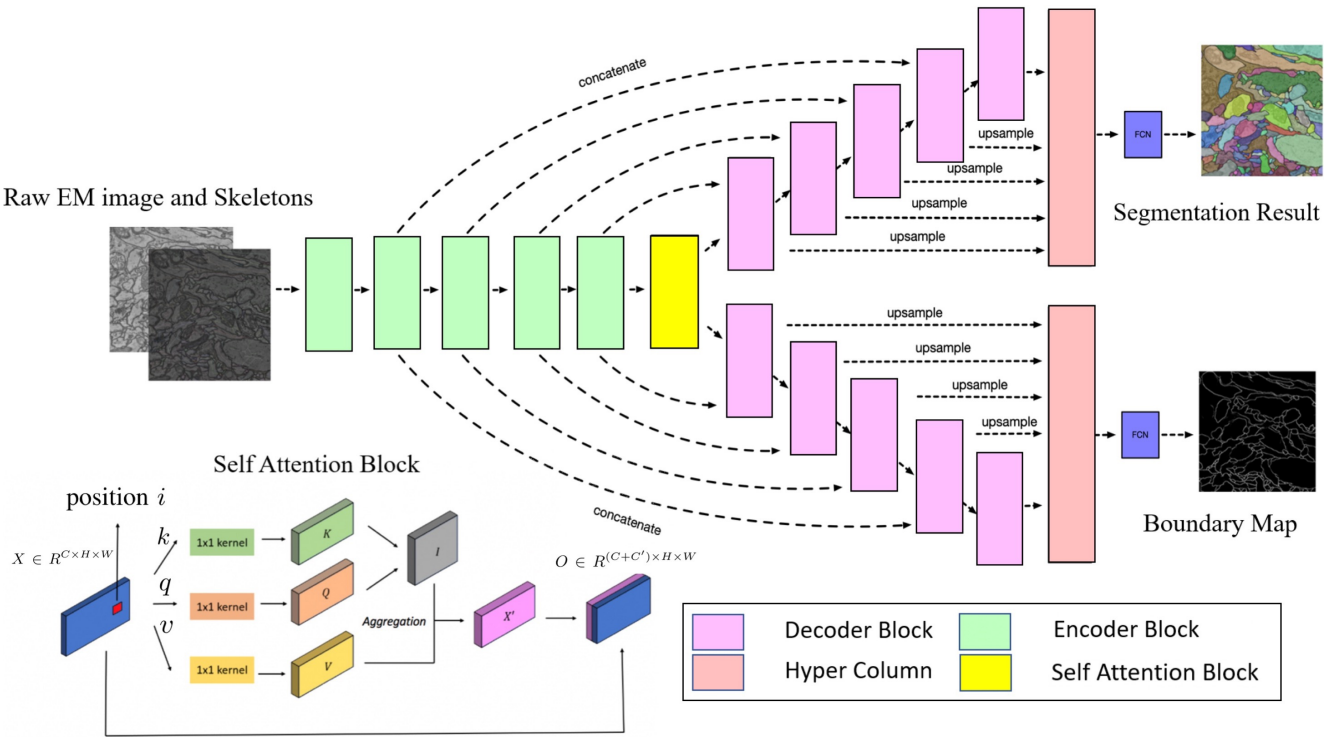

Figure 3. General network architecture of EMProof. An input of weak annotation is encoded with a UNET with self-attention and hypercolumn feature melting. A second branch encourages the network to produce crisp borders for k-class instance segmentation. A general view of the self-attention block is at the bottom left. The input (encoded features) is convolved with 1×1 kernels to produce key, query and value representations. The aggregated representation is concatenated with the input to form the attention's output.

existing methods. Our solution offers an improvement to existing highly accurate methods [17, 79] while being an end-to-end trainable without a dependency on affinity maps, hence defining a new state-of-the-art.

We also demonstrate the ability of EMProof to correct erroneous segmentation by using it to refine coarse segmentation in a super resolution settings. EMProof when used on low resolution images can accelerate manual reconstruction efforts in labs which lack computational power required to work with large datasets. We thus evaluate the predictive power of EMProof on the widely studied AC3 mouse cortex dataset (a large saturated human-made reconstruction) [41]. We show that EMProof can finely resolve the errors made by a low resolution reconstruction. EMProof accelerates the reconstruction from [41] by several folds while operating on coarse segmentations, down-sampled by a factor of 16x in each dimension (as shown in Fig. 5). EMProof manages to achieve an average 2-D VI score of 0.04.

We used standard metrics such as Variation of Information (VI) [50] and ARAND, which is the complement of Adjusted Rand index [68], for the experiments.

4.1. Implementation Details

Loss. As shown in Fig. 3, the network consists of two branches. The first branch predicts the multi class mask and the second predicts the boundary map. Cross-Entropy loss was used for the segmentation branch. For the boundary detection branch, several loss functions such as dice [53], lovsaz [8] and binary cross entropy loss were explored and evaluated. From our experiments we conclude that weighted sum of dice and cross entropy losses yields the best segmentation results. The weights set for dice and binary cross entropy losses were 0.25 and 0.75, respectively. The total loss for the network was defined as the weighted loss of the individual branches: $l_{\text{total}} = \lambda l_{\text{segmentation}} + (1-\lambda) l_{\text{border}}$, where λ is a hyperparameter used to balance the two branches outputs, $l_{\text{segmentation}}$ and l_{border} are the losses determined from the segmentation and boundary prediction branches, respectively.

Training Strategy. The networks were trained with the Stochastic Gradient Descent (SGD) algorithm with momentum of 0.9. The learning rate was set at 0.01 and was reduced in half every 7 epochs. The training batch size was 12. The networks were trained for 40 epochs on each dataset. For the first 10 epochs, the hyperparameter λ was

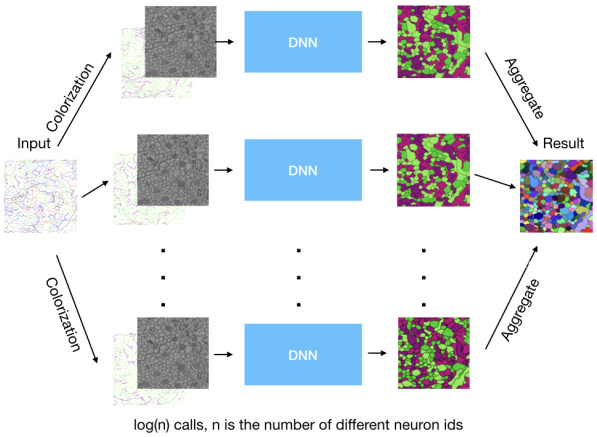

Figure 4. Overview of the internal workings of 3C within EMProof. First we colorize the labels by k different colors $l = \log(n)$ times where n is the number of objects in the image. This yields l different skeleton images, where each image now only has k colors. Each skeleton and EM image are then fed into the network. The model will then produce an expanded version of the skeleton images. All expanded skeleton images form together the final result (instance segmentation).

set to 0.4 and was subsequently reduced to 0.2, which was done in order to allow better gradient flow towards the segmentation arm at the start of training. Experiments were carried out on a Tesla V100 GPU with 32 GB of VRAM. The total training time was 4 hours per dataset.

4.2. Large-scale Reconstruction

We used two large-scale datasets to evaluate the scalability and speed of our technique: *Octopus vulgaris Brain* [11] and *Somatosensory Cortex* [41]. We demonstrate the ability of EMProof to correct sloppy annotations on both datasets. We assess the performance and speedup of EMProof on the octopus dataset by timing the reconstruction time of a sloppy versus meticulous annotation work. We also quantify the ability of EMProof to solve the super resolution task and produce a qualitatively meaningful reconstruction of complicated 3-D objects on the Somatosensory Cortex dataset.

Somatosensory Cortex. This dataset served in a pioneering connectomics study of the mouse somatosensory cortex [41]. EMProof improved the resolution of a segmentation that was down sampled by a factor of $2^4 = 16$ in each dimension. EMProof performed a difficult task since the EM images were also given at low resolution, making border detection highly challenging. The algorithm yet is able to give a fast and accurate overview of the shape of the complicated object as in Fig. 5. Reconstruction time took 20 minutes.

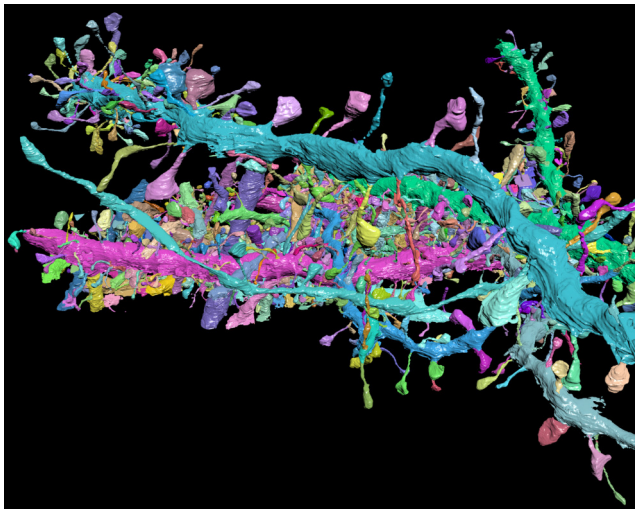

Figure 5. EMProof reconstructs dendrites from mouse cortex [41]. EMProof corrected annotations that were distorted by a 16x downsampling in each dimension.

Octopus vulgaris. The Octopus dataset is a 13K×7K pixels × 117 sections image-stack, 4 nm per pixel and 30 nm thick sections, collected from the vertical lobe of an adult *Octopus vulgaris*, shared with us for the purpose of this study by two leading labs which used our tool for connectomics reconstruction. For ground truth, high-quality and sloppy annotations were prepared in 8 hours and 1 hour, respectively. The training accuracy reached 96% and was qualitatively similar to the predictions on the test set.

4.3. Seeded Segmentation

EMProof relies on highly descriptive input of 2-D skeletons. To assess its predictive power, we trained EMProof on the more difficult task of expanding small round seeds. This allowed us to directly compare our technique to other works dealing with seeded segmentation. We use the challenging CREMI dataset [29] to evaluate our performance and benchmark it against standard pipelines mentioned below. CREMI is a dataset from MICCAI Challenge on Circuit Reconstruction from Electron Microscopy Images which is split into three datasets, each containing one image stack of 125 sections each from an adult *Drosophila melanogaster* brain tissue with the same EM image dimensions of 1250×1250. In order to ensure unbiased and fair benchmarking, we verify that our learning setup is identical to the ones previously published. Fig. 6 demonstrates the ability of EMProof to handle out-of-focus images and challenging geometries while managing to produce crisp and consistent boundary maps.

The following pipelines, which are also seeded approaches, are compared to our work. **Random Walker.** We slightly modified our network to directly predict bor-

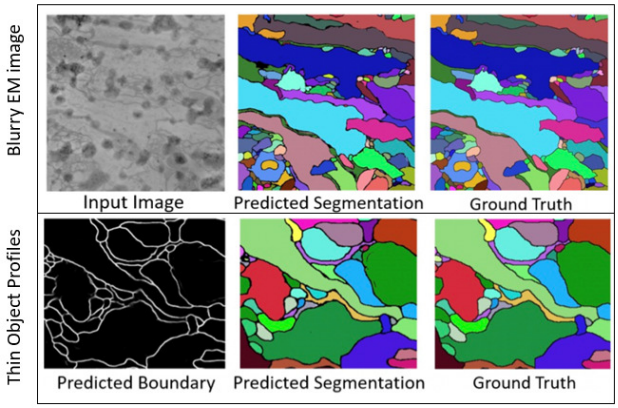

Figure 6. EMProof: Precise seeded segmentation of a blurry image and an image with thin objects from the CREMI dataset.

VI	WS	RW	Ours
CREMI A	0.053	0.142	**0.031**
CREMI B	0.189	0.283	**0.161**
CREMI C	0.211	0.433	**0.207**
Total	0.151	0.286	**0.133**
ARAND	WS	RW	Ours
CREMI A	0.013	0.033	**0.008**
CREMI B	0.043	0.149	**0.015**
CREMI C	0.045	0.228	**0.025**
Total	0.033	0.137	**0.016**

Table 1. Comparison between EMProof and the common baselines Watershed and Random Walker on the CREMI dataset. Metrics used are ARAND and VI.

der probability maps by removing the classification branch. After border prediction, we computed the instance segmentation result using the standard Random Walker algorithm given in [32]. **Watershed:** For the Watershed algorithm, we used the same predicted boundary probability and subsequently, we applied seeded watershed algorithm. **Learned Watershed:** We directly compared our results with the numbers published in [79]. **Learned Random Walker:** We compare our results with the numbers published in [17] while ensuring that same dataset and processing were used.

All experiments take round seeds as input for benchmarks in order to ensure as fair a comparison as possible. As shown in table 1, on all three dataset of CREMI, we largely exceed the result of watershed and random walker which used affinity graphs. This indicates that an end-to-end trained deep learning technique can outperform standard baselines for seeded segmentation.

We compare EMProof to the state-of-the-art in seeded segmentation in connectomics in the context of two experiments. In the first experiment, *one path*, we train our model to detect borders and then proceed with seeded watershed. We do not ablate any component of our architecture (squeeze-excitation, self-attention, hypervolume etc.). Our improved border predictions combined with standard region-growing algorithms outperforms strong baselines in connectomics (table 2). This suggests that seed-aware border prediction is key to the success of EMProof.

VI	LRW	LWS	Our: boundary	Our: boundary + segmentation
CREMI A	**0.0620**	-	-	0.068
CREMI B	0.193	-	-	**0.171**
CREMI C	0.232	-	0.214	**0.210**
Total	0.162	0.376	0.168	**0.150**
ARAND	LRW	LWS	Our: boundary	Our: boundary + segmentation
CREMI A	0.011	-	0.010	**0.007**
CREMI B	0.045	-	0.030	**0.025**
CREMI C	0.061	-	0.037	**0.033**
Total	0.039	0.082	0.026	**0.021**

Table 2. Comparison of Learned Random Walker, Learned Watershed, and EMProof on the seeded segmentation task on the public CREMI dataset. Metrics are adapted to those used in the previous state of the art. EMProof: In "Boundary," watershed is computed using the predicted boundaries; In "Boundary+segmentation," the 3C [51] results of EMProof are aggregated.

In the second experiment, we evaluate EMProof using both the border and segmentation outputs. Compared to the borders-only experiment above, EMProof indeed achieves further improvements in accuracy, although only marginally. This suggests that much of the success of EMProof is due to the seed-aware learning setup.

From table 1, we show quantitatively that EMProof outperformed all standard baselines when skeletons are taken as input, offering improvements across all three datasets. We also show from table 2 that EMProof is able to adapt to seeds that are not skeletons, while outperforming the state-of-the-art on two out of the three volumes.

4.4. Ablation Study

To further study the effectiveness of the EMProof architecture and the contribution of each of our design choices, we conducted extensive ablation experiments on the validation set of CREMI. Notably, as shown in table 3, a two-branch setting is essential for achieving top scores for seeded segmentation. Furthermore, the self-attention block helps to improve accuracy across the board, demonstrating its importance for EMProof.

VI	CREMI A	CREMI B	CREMI C
U-Net	0.113	0.679	0.558
U-Net + SCse [37]	0.104	0.559	0.428
U-Net + self-attention	0.068	0.419	0.347
U-Net + hypercolumn [34]	0.103	0.537	0.199
U-Net + two branch	0.041	0.316	**0.198**
U-Net + all (proposed)	**0.031**	**0.161**	0.207
ARAND	CREMI A	CREMI B	CREMI C
U-Net	0.033	0.057	0.658
U-Net + SCse	0.020	0.259	0.427
U-Net + self-attention	0.015	0.023	0.034
U-Net + hypercolumn	0.031	0.053	0.629
U-Net + two branch	0.013	0.021	0.032
U-Net + all(proposed)	**0.008**	**0.015**	**0.025**

Table 3. Comparison of different architectures by ablating EMProof. Metrics are computed on skeletal inputs and 3-pixel wide borders as in table 1.

4.5. Super Resolution

We assessed the robustness of EMProof by testing its ability to correct coarse segmentation that was produced on a low resolution image. To the best of our knowledge we are first to present a technique that allows annotators to work on low resolution images using machine learning to enhance the results of their segmentation. In this setting, a low resolution electron microscopy image is fed to EMProof alongside a coarse segmentation in order to correct the errors evident in the coarse segmentation. As with the other scenarios described in the paper, the coarse segmentation is skeletonized prior to the correction by EMProof.

We demonstrate this use case on the AC3 dataset which is an image stack used for benchmarking machine learning algorithms [41]. The AC3 dataset consists of 256 densely annotated images cropped from the larger 100,000 cubic microns mouse cortex volume by (∼80GB on disk) which was used in our large-scale experiments. Fig. 7 depicts predictions of high resolution segmentation from low resolution segmentation with a 16-folds improvement factor (in X and Y). The ability of our technique to super-resolve a segmentation from low resolution microscopy images, with negligible errors compared to ground truth, demonstrates resilience to poor image quality.

We quantified performance in this experiment by comparing the improvement in the segmentation quality before and after EMProof (Fig. 8). The improvement factor is defined as the ratio between VI score after correction to the VI score before correction. We observe that significant portion of the dataset obtain an improvement factor of 4-8x, with the majority achieving a VI under 0.04.

5. Conclusion

In this work, we demonstrate a simple end-to-end learning framework that outperforms the state-of-the-art of 2-D seeded segmentation of electron microscopy images, eliminating the dependency on affinity maps. We use this technique to accelerate the manual reconstruction of neurons from electron microscopy datasets. We believe our approach may apply to other reconstruction tasks in which an annotator provides a less accurate initial representation of an subject matter while relying on machine learning to improve and complete the task.

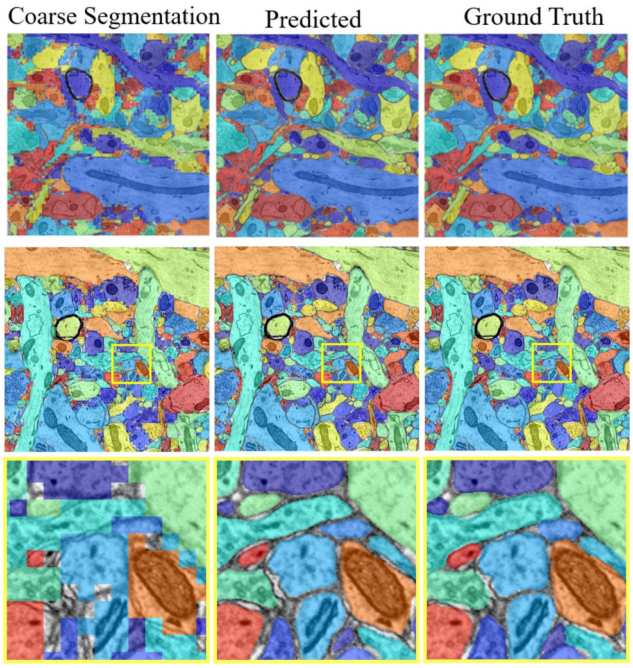

Figure 7. First two rows show examples of images from the AC3 dataset [41]. The last row magnifies a small region for better clarity. Columns show coarse segmentation inputs to EMProof, the super-resolved predictions of EMProof and ground truth.

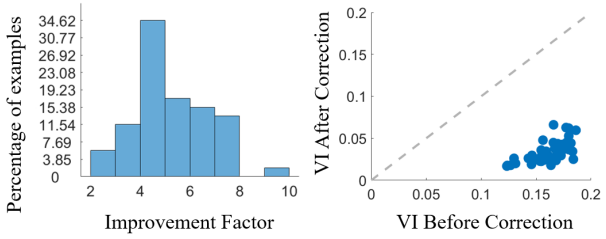

Figure 8. Comparative study of the impact of EMProof on "sloppy" annotations. The left figure is Histogram of Improvement Factor and the right is a comparison of VI before and after correction

6. Acknowledgements

This work was supported by a grant from Thermo Fisher Scientific.

References

[1] Larry F Abbott, Davi D Bock, Edward M Callaway, Winfried Denk, Catherine Dulac, Adrienne L Fairhall, Ila Fiete, Kristen M Harris, Moritz Helmstaedter, Viren Jain, et al. The mind of a mouse. *Cell*, 182(6):1372–1376, 2020. 1

[2] Ignacio Arganda-Carreras, Srinivas C Turaga, Daniel R Berger, Dan Cireşan, Alessandro Giusti, Luca M Gambardella, Jürgen Schmidhuber, Dmitry Laptev, Sarvesh Dwivedi, Joachim M Buhmann, et al. Crowdsourcing the creation of image segmentation algorithms for connectomics. *Frontiers in neuroanatomy*, 9:142, 2015. 4

[3] Mária Ashaber, Yusuke Tomina, Pegah Kassraian, Eric A Bushong, William B Kristan Jnr, Mark H Ellisman, and Daniel A Wagenaar. Anatomy and activity patterns in a multifunctional motor neuron and its surrounding circuits. *Elife*, 10:e61881, 2021. 2

[4] Alberto Bailoni, Constantin Pape, Nathan Hütsch, Steffen Wolf, Thorsten Beier, Anna Kreshuk, and Fred A. Hamprecht. Gasp, a generalized framework for agglomerative clustering of signed graphs and its application to instance segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11645–11655, June 2022. 1

[5] Alberto Bailoni, Constantin Pape, Steffen Wolf, Anna Kreshuk, and Fred A. Hamprecht. Proposal-free volumetric instance segmentation from latent single-instance masks, 2020. 1

[6] Thorsten Beier, Constantin Pape, Nasim Rahaman, Timo Prange, Stuart Berg, Davi D Bock, Albert Cardona, Graham W Knott, Stephen M Plaza, Louis K Scheffer, et al. Multicut brings automated neurite segmentation closer to human performance. *Nature Methods*, 14(2):101–102, 2017. 1, 3

[7] Stuart Berg, Dominik Kutra, Thorben Kroeger, Christoph N Straehle, Bernhard X Kausler, Carsten Haubold, Martin Schiegg, Janez Ales, Thorsten Beier, Markus Rudy, et al. Ilastik: interactive machine learning for (bio) image analysis. *Nature Methods*, pages 1–7, 2019. 2

[8] Maxim Berman, Amal Rannen Triki, and Matthew B. Blaschko. The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks, 2017. 5

[9] Gedas Bertasius, Lorenzo Torresani, Stella X Yu, and Jianbo Shi. Convolutional random walk networks for semantic image segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 858–866, 2017. 3

[10] Johanna Beyer, Jakob Troidl, Saeed Boorboor, Markus Hadwiger, Arie Kaufman, and Hanspeter Pfister. A survey of visualization and analysis in high-resolution connectomics. In *Computer Graphics Forum*, volume 41, pages 573–607. Wiley Online Library, 2022. 1

[11] Flavie Bidel, Yaron Meirovitch, Richard Lee Schalek, Xiaotang Lu, Elisa Catherine Pavarino, Fuming Yang, Adi Peleg, Yuelong Wu, Tal Shomrat, Daniel Raimund Berger, et al. Connectomics of the octopus vulgaris vertical lobe provides insight into conserved and novel principles of a memory acquisition network. *bioRxiv*, pages 2022–10, 2022. 2, 6

[12] Yuri Boykov and Gareth Funka-Lea. Graph cuts and efficient nd image segmentation. *International journal of computer vision*, 70(2):109–131, 2006. 3

[13] Philippe Burlina. Mrcnn: A stateful fast r-cnn. In *2016 23rd International conference on pattern recognition (ICPR)*, pages 3518–3523. IEEE, 2016. 1

[14] Yue Cao, Jiarui Xu, Stephen Lin, Fangyun Wei, and Han Hu. Gcnet: Non-local networks meet squeeze-excitation networks and beyond. *arXiv preprint arXiv:1904.11492*, 2019. 1

[15] Wallace Casaca, Luis Gustavo Nonato, and Gabriel Taubin. Laplacian coordinates for seeded image segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 384–391, 2014. 2

[16] Brendan Celii, Stelios Papadopoulos, Zhuokun Ding, Paul G Fahey, Eric Wang, Christos Papadopoulos, Alexander Kunin, Saumil Patel, J Alexander Bae, Agnes L Bodor, et al. Neurd: A mesh decomposition framework for automated proofreading and morphological analysis of neuronal em reconstructions. *bioRxiv*, pages 2023–03, 2023. 1

[17] Lorenzo Cerrone, Alexander Zeilmann, and Fred A Hamprecht. End-to-end learned random walker for seeded image segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 12559–12568, 2019. 2, 3, 5, 7

[18] D. J. Chen, H. T. Chen, and L. W. Chang. Swipecut: Interactive segmentation via seed grouping. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(9):2959–2970, 2020. 3

[19] Kai Chen, Jiangmiao Pang, Jiaqi Wang, Yu Xiong, Xiaoxiao Li, Shuyang Sun, Wansen Feng, Ziwei Liu, Jianping Shi, Wanli Ouyang, Chen Change Loy, and Dahua Lin. Hybrid task cascade for instance segmentation, 2019. 1

[20] Liang Chen, Paul Bentley, Kensaku Mori, Kazunari Misawa, Michitaka Fujiwara, and Daniel Rueckert. Drinet for medical image segmentation. *IEEE transactions on medical imaging*, 37(11):2453–2462, 2018. 1

[21] Dan Ciresan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems*, pages 2843–2851, 2012. 3

[22] Camille Couprie, Pauline Luc, and Jakob Verbeek. Joint future semantic and instance segmentation prediction. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, September 2018. 3

[23] Jean Cousty, Gilles Bertrand, Laurent Najman, and Michel Couprie. Watershed cuts: Thinnings, shortest path forests, and topological watersheds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5):925–939, 2009. 3

[24] Antonio Criminisi, Jamie Shotton, and Ender Konukoglu. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trendstextregistered in Computer Graphics and Vision*, 7:81–227, 01 2011. 3

[25] Konstantin Dimitriev, T. Parag, Brian Matejek, A. Kaufman, and H. Pfister. Efficient correction for em connectomics with skeletal representation. In *BMVC*, 2018. 2

[26] Michael Donoser, Martin Urschler, Martin Hirzer, and Horst Bischof. Saliency driven total variation segmentation. In *2009 IEEE 12th International Conference on Computer Vision*, pages 817–824. IEEE, 2009. 3

[27] Sven Dorkenwald, Claire E McKellar, Thomas Macrina, Nico Kemnitz, Kisuk Lee, Ran Lu, Jingpeng Wu, Sergiy Popovych, Eric Mitchell, Barak Nehoran, et al. Flywire: online community for whole-brain connectomics. *Nature methods*, 19(1):119–128, 2022. 1

[28] Michal Drozdzal, Gabriel Chartrand, Eugene Vorontsov, Mahsa Shakeri, Lisa Di Jorio, An Tang, Adriana Romero, Yoshua Bengio, Chris Pal, and Samuel Kadoury. Learning normalized inputs for iterative estimation in medical image segmentation. *Medical image analysis*, 44:1–13, 2018. 1

[29] Jan Funke, Stephan Saalfeld, Davi D Bock, Srinivas C Turaga, and Erik Perlman. *CREMI*, 2016 (accessed: 2019-05-15). 2, 6

[30] Jan Funke, Fabian Tschopp, William Grisaitis, Arlo Sheridan, Chandan Singh, Stephan Saalfeld, and Srinivas C Turaga. Large scale image segmentation with structured loss based deep learning for connectome reconstruction. *IEEE transactions on pattern analysis and machine intelligence*, 41(7):1669–1680, 2018. 1

[31] Octavio Gómez, Jesús A. González, and Eduardo F. Morales. Image segmentation using automatic seeded region growing and instance-based learning. In Luis Rueda, Domingo Mery, and Josef Kittler, editors, *Progress in Pattern Recognition, Image Analysis and Applications*, pages 192–201, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. 3

[32] Leo Grady. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (11):1768–1783, 2006. 3, 7

[33] Daniel Haehn, Verena Kaynig, James Tompkin, Jeff W. Lichtman, and Hanspeter Pfister. Guided proofreading of automatic segmentations for connectomics. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 1, 2

[34] Bharath Hariharan, Pablo Arbelaez, Ross Girshick, and Jitendra Malik. Hypercolumns for object segmentation and fine-grained localization. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2015. 8

[35] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017. 1

[36] Moritz Helmstaedter, Kevin L Briggman, Srinivas C Turaga, Viren Jain, H Sebastian Seung, and Winfried Denk. Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*, 500(7461):168, 2013. 2

[37] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018. 3, 4, 8

[38] Jeremy Jancsary, Sebastian Nowozin, and Carsten Rother. Loss-specific training of non-parametric image restoration models: A new state of the art. In *European Conference on Computer Vision*, pages 112–125. Springer, 2012. 3

[39] Michał Januszewski, Jörgen Kornfeld, Peter H Li, Art Pope, Tim Blakely, Larry Lindsey, Jeremy Maitin-Shepard, Mike Tyka, Winfried Denk, and Viren Jain. High-precision automated reconstruction of neurons with flood-filling networks. *Nature methods*, 15(8):605, 2018. 1, 2, 3

[40] Neha Karlupia, Richard L Schalek, Yuelong Wu, Yaron Meirovitch, Donglai Wei, Alexander W Charney, Brian H Kopell, and Jeff W Lichtman. Immersion fixation and staining of multicubic millimeter volumes for electron microscopy–based connectomics of human brain biopsies. *Biological Psychiatry*, 2023. 2

[41] Narayanan Kasthuri, Kenneth Jeffrey Hayworth, Daniel Raimund Berger, Richard Lee Schalek, José Angel Conchello, Seymour Knowles-Barley, Dongil Lee, Amelio Vázquez-Reina, Verena Kaynig, Thouis Raymond Jones, et al. Saturated reconstruction of a volume of neocortex. *Cell*, 162(3):648–661, 2015. 2, 3, 5, 6, 8

[42] Verena Kaynig, Amelio Vazquez-Reina, Seymour Knowles-Barley, Mike Roberts, Thouis R Jones, Narayanan Kasthuri, Eric Miller, Jeff Lichtman, and Hanspeter Pfister. Large-scale automatic reconstruction of neuronal processes from electron microscopy images. *Medical image analysis*, 22(1):77–88, 2015. 3

[43] Margret Keuper, Siyu Tang, Bjorn Andres, Thomas Brox, and Bernt Schiele. Motion segmentation & multiple object tracking by correlation co-clustering. *IEEE transactions on pattern analysis and machine intelligence*, 2018. 1

[44] Kisuk Lee, Jonathan Zung, Peter Li, Viren Jain, and H Sebastian Seung. Superhuman accuracy on the SNEMI3D connectomics challenge. *arXiv preprint arXiv:1706.00120*, 2017. 1, 3

[45] Justin Liang, Namdar Homayounfar, Wei-Chiu Ma, Yuwen Xiong, Rui Hu, and Raquel Urtasun. Polytransform: Deep polygon transformer for instance segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 3

[46] Jeff W Lichtman, Hanspeter Pfister, and Nir Shavit. The big data challenges of connectomics. *Nature neuroscience*, 17(11):1448–1454, 2014. 1

[47] Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia. Path aggregation network for instance segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8759–8768, 2018. 1

[48] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015. 4

[49] Brian Matejek, Daniel Haehn, Haidong Zhu, Donglai Wei, Toufiq Parag, and Hanspeter Pfister. Biologically-constrained graphs for global connectomics reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019. 2

[50] Marina Meilă. Comparing clusterings. *Journal of multivariate analysis*, 98(5):873–895, 2007. 5

[51] Yaron Meirovitch, Lu Mi, Hayk Saribekyan, Alexander Matveev, David Rolnick, and Nir Shavit. Cross-classification

clustering: An efficient multi-object tracking technique for 3-d instance segmentation in connectomics, 2019. 1, 2, 3, 4, 7

[52] F. Meng, H. Li, Q. Wu, K. N. Ngan, and J. Cai. Seeds-based part segmentation by seeds propagation and region convexity decomposition. *IEEE Transactions on Multimedia*, 20(2):310–322, 2018. 3

[53] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation, 2016. 5

[54] Josh Lyskowski Morgan, Daniel Raimund Berger, Arthur Willis Wetzel, and Jeff William Lichtman. The fuzzy logic of network connectivity in mouse visual thalamus. *Cell*, 165(1):192–206, 2016. 2

[55] Alessandro Motta, Manuel Berning, Kevin M Boergens, Benedikt Staffler, Marcel Beining, Sahil Loomba, Philipp Hennig, Heiko Wissler, and Moritz Helmstaedter. Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science*, page eaay3134, 2019. 2

[56] S Nagaraju, Manish Kashyap, Sandeep Kumar, and Mahua Bhattacharya. Affinity based seeded region growing algorithm for medical image segmentation. In *2015 International Conference on Computing and Network Communications (CoCoNet)*, pages 725–730. IEEE, 2015. 2

[57] C. Pape, T. Beier, P. Li, V. Jain, D. D. Bock, and A. Kreshuk. Solving large multicut problems for connectomics via domain decomposition. In *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, pages 1–10, 2017. 2

[58] Toufiq Parag, Fabian Tschopp, William Grisaitis, Srinivas C Turaga, Xuewen Zhang, Brian Matejek, Lee Kamentsky, Jeff W Lichtman, and Hanspeter Pfister. Anisotropic EM segmentation by 3D affinity learning and agglomeration. *arXiv preprint arXiv:1707.08935*, 2017. 3

[59] S. Park, H. S. Lee, and J. Kim. Seed growing for interactive image segmentation with geodesic voting. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 2564–2568, 2016. 3

[60] Chanchan Qin, Guoping Zhang, Yicong Zhou, Wenbing Tao, and Zhiguo Cao. Integration of the saliency-based seed extraction and random walks for image segmentation. *Neurocomputing*, 129:378–391, 2014. 3

[61] David Rolnick, Yaron Meirovitch, Toufiq Parag, Hanspeter Pfister, Viren Jain, Jeff W Lichtman, Edward S Boyden, and Nir Shavit. Morphological error detection in 3d segmentations. *arXiv preprint arXiv:1705.10882*, 2017. 2

[62] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, pages 234–241. Springer, 2015. 3, 4

[63] Kegan GG Samuel and Marshall F Tappen. Learning optimized map estimates in continuously-valued mrf models. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 477–484. IEEE, 2009. 3

[64] Philipp J Schubert, Sven Dorkenwald, Michał Januszewski, Viren Jain, and Joergen Kornfeld. Learning cellular morphology with neural networks. *Nature Communications*, 10, June 2019. 2

[65] Ali Kemal Sinop and Leo Grady. A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm. In *ICCV*, volume 1, page 3. Citeseer, 2007. 2

[66] Zhiyu Tan, Liang Wan, Wei Feng, and Chi-Man Pun. Image co-saliency detection by propagating superpixel affinities. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2114–2118. IEEE, 2013. 3

[67] Humera Tariq, Tahseen Jilani, Usman Amjad, and SM Aqil Burney. Novel seed selection and conceptual region growing framework for medical image segmentation. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 10(1):6–19, 2019. 1

[68] Ranjith Unnikrishnan, Caroline Pantofaru, and Martial Hebert. Toward objective evaluation of image segmentation algorithms. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6):929–944, 2007. 5

[69] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017. 3, 4

[70] Csaba Verasztó, Sanja Jasek, Martin Gühmann, Réza Shahidi, Nobuo Ueda, James David Beard, Sara Mendes, Konrad Heinz, Luis Alberto Bezares-Calderón, Elizabeth Williams, and Gáspár Jékely. Whole-animal connectome and cell-type complement of the three-segmented platynereis dumerilii larva. *bioRxiv*, 2020. 2

[71] Paul Voigtlaender, Michael Krause, Aljosa Osep, Jonathon Luiten, Berin Balachandar Gnana Sekar, Andreas Geiger, and Bastian Leibe. Mots: Multi-object tracking and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7942–7951, 2019. 1

[72] Guotai Wang, Wenqi Li, Maria A Zuluaga, Rosalind Pratt, Premal A Patel, Michael Aertsen, Tom Doel, Anna L David, Jan Deprest, Sébastien Ourselin, et al. Interactive medical image segmentation using deep learning with image-specific fine tuning. *IEEE transactions on medical imaging*, 37(7):1562–1573, 2018. 1

[73] Puyang Wang, Nick G Cuccolo, Rachana Tyagi, Ilker Hacihaliloglu, and Vishal M Patel. Automatic real-time cnn-based neonatal brain ventricles segmentation. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 716–719. IEEE, 2018. 1

[74] Qiang Wang, Li Zhang, Luca Bertinetto, Weiming Hu, and Philip HS Torr. Fast online object tracking and segmentation: A unifying approach. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1328–1338, 2019. 1

[75] John G White, Eileen Southgate, J Nichol Thomson, and Sydney Brenner. The structure of the nervous system of the nematode caenorhabditis elegans. *Philos Trans R Soc Lond B Biol Sci*, 314(1165):1–340, 1986. 1

[76] Daniel Witvliet, Ben Mulcahy, James K Mitchell, Yaron Meirovitch, Daniel K Berger, Yuelong Wu, Yufang Liu, Wan Xian Koh, Rajeev Parvathala, Douglas Holmyard, et al. Connectomes across development reveal principles of brain maturation in c. elegans. *bioRxiv*, 2020. 2

[77] Steffen Wolf, Yuyan Li, Constantin Pape, Alberto Bailoni, Anna Kreshuk, and Fred A. Hamprecht. The semantic mutex watershed for efficient bottom-up semantic instance segmentation. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision – ECCV 2020*, pages 208–224, Cham, 2020. Springer International Publishing. 3

[78] Steffen Wolf, Constantin Pape, Alberto Bailoni, Nasim Rahaman, Anna Kreshuk, Ullrich Kothe, and FredA Hamprecht. The mutex watershed: efficient, parameter-free image partitioning. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 546–562, 2018. 3

[79] Steffen Wolf, Lukas Schott, Ullrich Kothe, and Fred Hamprecht. Learned watershed: End-to-end learning of seeded segmentation. *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct 2017. 2, 3, 5, 7

[80] Daniel Xenes, Lindsey M Kitchell, Patricia K Rivlin, Rachel Brodsky, Hannah Gooden, Justin Joyce, Diego Luna, Raphael Norman-Tenazas, Devin Ramsden, Kevin Romero, et al. Neuvue: A framework and workflows for high-throughput electron microscopy connectomics proofreading. *bioRxiv*, pages 2022–07, 2022. 1

[81] Hang Zhang, Chongruo Wu, Zhongyue Zhang, Yi Zhu, Zhi Zhang, Haibin Lin, Yue Sun, Tong He, Jonas Mueller, R. Manmatha, Mu Li, and Alexander Smola. Resnest: Split-attention networks, 2020. 3

[82] Hengshuang Zhao, Yi Zhang, Shu Liu, Jianping Shi, Chen Change Loy, Dahua Lin, and Jiaya Jia. Psanet: Point-wise spatial attention network for scene parsing. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 267–283, 2018. 3, 4

[83] Xiaojin Zhu, Zoubin Ghahramani, and John D Lafferty. Semi-supervised learning using gaussian fields and harmonic functions. In *Proceedings of the 20th International conference on Machine learning (ICML-03)*, pages 912–919, 2003. 3

[84] X. Zhu, C. C. Loy, and S. Gong. Video synopsis by heterogeneous multi-source correlation. In *2013 IEEE International Conference on Computer Vision*, pages 81–88, 2013. 3