ACCURATE AND EFFICIENT SOMA RECONSTRUCTION IN A FULL ADULT FLY BRAIN

Anonymous authors
Paper under double-blind review

ABSTRACT

Neuron reconstruction in a full adult fly brain from high-resolution electron microscopy (EM) data is regarded as a cornerstone for neuroscientists to explore how neurons inspire intelligence. As the central part of neurons, somas in the full brain indicate the origin of neurogenesis and neural functions. However, due to the absence of EM datasets specifically annotated for somas, existing deep learning-based neuron reconstruction methods cannot directly provide accurate soma distribution and morphology. Moreover, full brain neuron reconstruction remains extremely time-consuming due to the unprecedentedly large size of EM data. In this paper, we develop an efficient soma reconstruction method for obtaining accurate soma distribution and morphology information in a full adult fly brain. To this end, we first make a high-resolution EM dataset with fine-grained 3D manual annotations on somas. Relying on this dataset, we propose an efficient, two-stage deep learning algorithm for predicting accurate locations and boundaries of 3D soma instances. Further, we deploy a parallelized, high-throughput data processing pipeline for executing the above algorithm on the full brain. Finally, we provide quantitative and qualitative results to validate the superiority of the proposed method, as well as comprehensive statistics of the reconstructed somas in the full adult fly brain from the biological perspective.

1 INTRODUCTION

Drosophila melanogaster, also named as the fruit fly, is an organism with intelligent behaviors including perception, learning, and judgment (Ofstad et al., 2011; Dickinson & Muijres, 2016; Owald & Waddell, 2015). It has a complete and relatively simple neural system (Pavlou & Goodwin, 2013; Burne et al., 2011). The interactions among neurons in the system guide the drosophila's intelligent behaviors (Ohyama et al., 2015; Januszewski et al., 2018; Cachero et al., 2010; Gerstein, 1960). Therefore, the study of drosophila neurons, which has fascinated neuroscientists for more than a century (Simpson, 2009; Gruntman & Turner, 2013; Busch et al., 2009; van Naters & Carlson, 2007; Bodmer & Jan, 1987), has key implications for understanding how the brains of living organisms produce intelligence (Armstrong & van Hemert, 2009; Takemura et al., 2017).

As the central part of the neuron, the soma maintains the neuron structure and controls the formation of neurites (Luengo-Sanchez et al., 2015; Ito et al., 2013). Studies have shown that the location and morphology of somas in the full brain are related to neural development and the neural logic function (Hartenstein, 2011; Bae et al., 2021), and the number of somas is related to the complexity of the brain and the age of the living body (Andersen et al., 2003; Lent et al., 2012). Therefore, it is of great biological significance to investigate soma reconstruction in the full brain of model organisms such as drosophila.

Traditional studies in this field are mainly based on brain images collected by optical microscopies (Rein et al., 2002; García-Cabezas et al., 2016). The soma structure is first stained with specific staining proteins, and the confocal images collected could show fluorescence staining signals, so as to obtain the soma distribution of the full brain of drosophila. However, the resolution of confocal images is low, making it difficult to obtain the exact morphology of each soma. Based on the assumption that each cell has only one nucleus, the isotropic fractionator method (Godfrey et al., 2021; Herculano-Houzel & Lent, 2005) is used to obtain the number of cell bodies in the full brain

of drosophila. This method destroys the brain structure during the production of cell suspension, so the distribution of somas in the full brain cannot be obtained.

Recently, with the development of high-speed electron microscopy (EM) scanning technology, high-resolution EM image datasets of different species including drosophila (Takemura et al., 2015), mouse (Motta et al., 2019), and human (Shapson-Coe et al., 2021) have been successfully acquired, and the full adult fly brain (FAFB) dataset (Zheng et al., 2018) imaged from a complete drosophila brain can be regarded as a representative. Based on these datasets, advanced deep learning algorithms are developed to automatically reconstruct neurons (Januszewski et al., 2018; Funke et al., 2018) and cell nucleus (Mu et al., 2021) in EM images for connectomics study. Meanwhile, parallel and distributed data processing pipelines (Wu et al., 2021; Shapson-Coe et al., 2021) are proposed to deploy these algorithms on large-scale EM datasets. However, due to the lack of high-resolution EM datasets specifically annotated for somas, existing works cannot directly provide accurate soma distribution and morphology information.

In this paper, we make one of the first efforts to develop an efficient soma reconstruction method in a full adult fly brain, aiming to obtain accurate soma distribution and morphology information for this model organism. The contributions of this work are four aspects:

1. We make a high-resolution EM soma dataset with fine-grained 3D manual annotations for more than 8×10^9 voxels. To the best of our knowledge, this dataset is the first of its kind.
2. Relying on the above dataset, we propose an efficient, two-stage deep learning algorithm for soma instance segmentation, which first localizes somas and then generates their 3D boundaries with high fidelity.
3. We deploy a parallelized, high-throughput data processing pipeline for executing our algorithm on the full brain, fulfilling the soma reconstruction task on a 90-GPU cluster within 4 days.
4. We provide quantitative and qualitative results for evaluating the accuracy and efficiency of the proposed method, along with comprehensive statistics of the reconstructed somas in the full adult fly brain, including count, size, distribution and morphology.

We believe our work will contribute to the study of the drosophila neural system. The dataset will be released to facilitate future research along this line. Code and a 4K video of the full brain reconstruction result are now available through the links provided in the Appendix.

2 Related Work

Neuron Reconstruction. Recent works of neuron reconstruction in EM images, such as FFN (Januszewski et al., 2018) and MALA (Funke et al., 2018), embrace the power of deep learning and obtain neuron instance information based on connectivity, pre-labeling, etc. However, due to the absence of elaborate annotations and specific designs for somas in these algorithms, the shape and position of the soma cannot be accurately predicted. Moreover, these algorithms are generally time-consuming when dealing with extremely large-scale EM data. For example, MALA has complex post-processing procedures based on optimization of traditional methods, and FFN has a very large number of sliding windows during inference, making them inefficient for fast soma reconstruction in the full brain of drosophila. More recent works (Li et al., 2019; Dorkenwald et al., 2022) reconstruct all neurons on the FAFB dataset. However, one cannot directly obtain the accurate distribution and morphology information of somas based on these results, since each generated segment often contains all parts of a fly neuron: a soma, dendrites, axon terminals, and a primary neurite. Therefore, our work on independent soma reconstruction is complementary to neuron reconstruction for connectomics study.

Nucleus Reconstruction. As a parallel line, there are two recent works on nucleus reconstruction in EM images. Although the soma has a one-to-one correspondence to the nucleus, the nucleus is often located in the center of the soma with a regular spherical shape, which cannot reflect the morphology of the soma. Mu et al. (Mu et al., 2021) reconstruct all cell nuclei on the FAFB dataset by using a standard 2D U-Net to predict the binary classification of individual pixels as either nucleus or not. However, this method lacks the utilization of 3D structural information and is difficult to

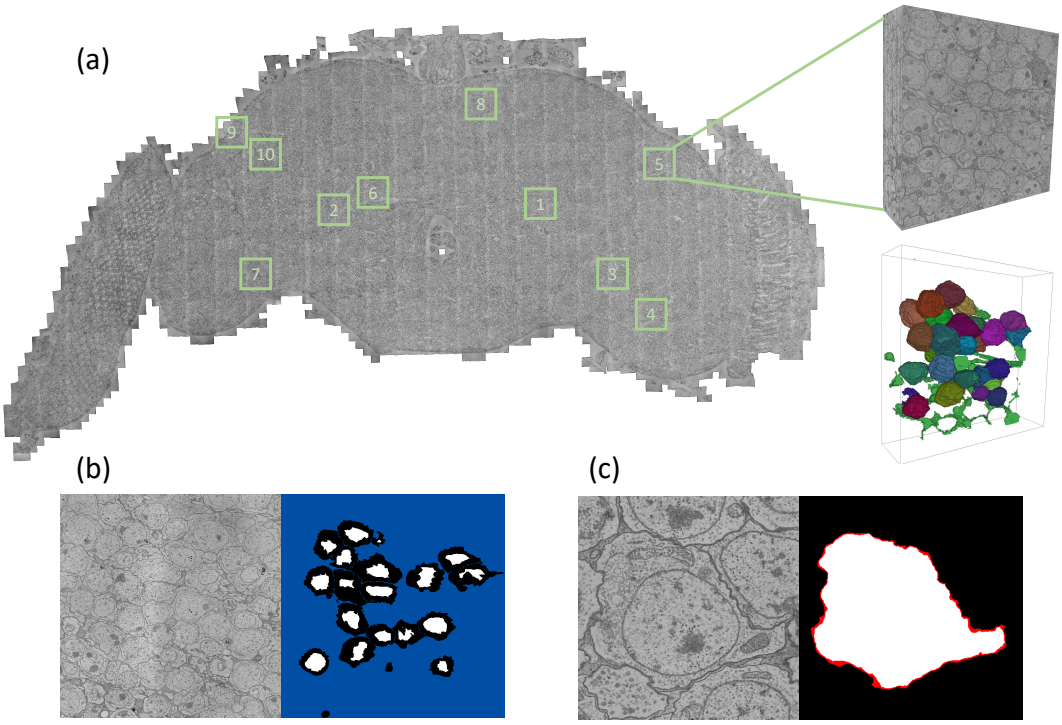

Figure 1: (a) An overview of the FAFB dataset. We select 10 EM blocks from the full brain, and annotate them to make our label sets. (b) A visualization example for our localization label set. The white and black colors indicate the seed and non-seed annotations, respectively. The area in blue color is unannotated. (c) A visualization example of the segmentation label set. The white and black colors indicate the soma-of-interest and background annotations, respectively. The red color indicates the boundary annotation of the soma-of-interest. For simplicity and clarity, we visualize (b) as a slice of the 3D block and (c) as a slice of the 3D patch.

apply to densely distributed somas with complex shapes. Lin et al. (Lin et al., 2021) make a neuronal nuclei instance segmentation EM dataset at the sub-cubic millimeter scale and propose a hybrid representation segmentation method by directly adopting a 3D model to predict multiple complete objects simultaneously. However, this method is not suitable for high-resolution EM images at the nanometer scale, since somas with much larger sizes cannot be processed by the model due to the limited GPU memory. Different from existing nucleus segmentation methods, our proposed soma segmentation method not only adopts a 3D model to consider the 3D structure of complete somas but also can directly process high-resolution EM images with an elaborate two-stage instance segmentation algorithm.

3 EM Soma Dataset with Fine-Grained Annotation

We make an EM adult drosophila soma (EMADS) dataset with fine-grained manual annotations. EMADS contains 204 completely annotated 3D somas with different sizes and morphologies derived from 10 apart regions in a full adult fly brain.

Source and Preparation. The EM images of the adult drosophila brain that we annotate originate from FAFB (Zheng et al., 2018). FAFB is the world's first EM dataset for a complete drosophila brain. FAFB is imaged at the synaptic resolution and has been processed with stitching and alignment. There are a total of $7,062$ sections in FAFB, and the full resolution of each section is $286,720 \times 155,648$ which is partitioned into $8,192 \times 8,192$ images, resulting in 40 TB data in storage. Thanks to the high-resolution imaging technique, biological structures and boundaries in FAFB are clear, making fine-grained manual annotations on the voxel-level data possible. However,

the annotation workload on such high-resolution, large-scale data is heavy. Before annotation, we first downsample FAFB at both the x-axis and the y-axis with a factor of 4. The physical resolution of the downsampled dataset is 16nm $\times$ 16nm $\times$ 40nm ($x \times y \times z$). In this way, the annotation workload for each soma is lessened so that we can annotate more somas in diverse sizes and morphologies. Besides, the reduced resolution also lessens the computational burden in our method and makes it affordable for 3D deep networks. We store the full brain data in our defined 3D image blocks, and the size of each block is $1,836 \times 1,836 \times 186$. The 3D image blocks are arranged in the order of x, y (in the lateral direction) and z (in the axial direction) axis. By loading these blocks in the axial orders continuously, we can obtain the full brain data. The parallel data processing pipeline in our method is detailed in Sec. 5.

Selective Annotations. Many previous works (Chiang et al., 2011; Ito et al., 2014) utilize specific biological stains to mark somas from 3D confocal images to indicate a general soma distribution in the full drosophila brain. Depending on this distribution, we select 10 apart image blocks from FAFB according to the location and density of somas in the full brain. We then organize a group of human annotators to annotate a part of somas and background areas inside these 10 image blocks, with an efficient and semi-automatic annotation tool VAST (Berger et al., 2018). The annotations are conducted at the voxel level, with the boundary of somas to be annotated precisely. 20 master and Ph.D. students majoring in neuroscience or computer vision devote themselves to this annotation task. We annotate 204 complete somas with different sizes and morphologies inside these 10 image blocks, and each soma takes more than two hours for one person to annotate. All 10 blocks are used to ensure the diversity of soma distribution and morphology. At the same time, weighing the workload of the annotators and the amount of data to be annotated, we only annotate about 20 somas in each block, and there are still unlabeled areas. The total number of annotated voxels is more than 8×10^9. The average size of these somas is around $300 \times 300 \times 100$ $voxel^3$. The annotations are instance-wise, which means each soma has a unique instance number that is different from other somas. After the group annotation, two experts check every annotated soma in a cross manner, and correct the annotation mistakes. Finally, we obtain 10 EM blocks with 10 corresponding label blocks, each with a resolution of $1,836 \times 1,836 \times 186$. We visualize one of these blocks as an example in Figure 1(a).

Localization Label Set. We first employ our annotations to make a label set for localizing somas in a given EM block. We perform a binary morphological erosion operation on the annotation of each soma in its corresponding label block. Through this operation, the eroded annotations of somas are labeled as instance seeds, which provides location information for each annotated soma. In addition, since the erosion operation preserves the morphology of each soma to a certain extent, we can obtain the rough size of the soma according to the size of the corresponding instance seed. After that, each labeled block in the localization label set consists of two types of labels besides the unannotated areas. These two labels denote the instance seeds of the annotated somas and the non-seed areas (which can be either the erosion areas or the non-soma areas). An example is given in Figure 1(b). This localization label set, stored in blocks with a uniform size after further downsampling, is used to localize somas in an efficient way in the first stage of our segmentation algorithm (see details in Sec. 4).

Segmentation Label Set. We then employ our partially annotated blocks to make the other label set for segmenting somas from patches. Specifically, we crop the annotated somas from their blocks to patches. The size of each patch is slightly larger than the size of the soma, and each soma is at the center of its patch. Since many somas are next to each other and thus it is hard to distinguish instances, we devise the segmentation label set with three types of labels to tackle this problem. We erode the annotation of each soma inside its patch by two voxels. We then label the eroded soma at the center of the patch as the soma-of-interest, and label the erosion area as the soma boundary. The rest area in a patch except for the soma-of-interest and the soma boundary is regarded as a separate label. Therefore, the segmentation label set consists of three types of segmentation labels. An example is given in Figure 1(c). This segmentation label set, stored in patches with variable sizes, is used to segment somas in an accurate way in the second stage of our segmentation algorithm (see details in Sec. 4).

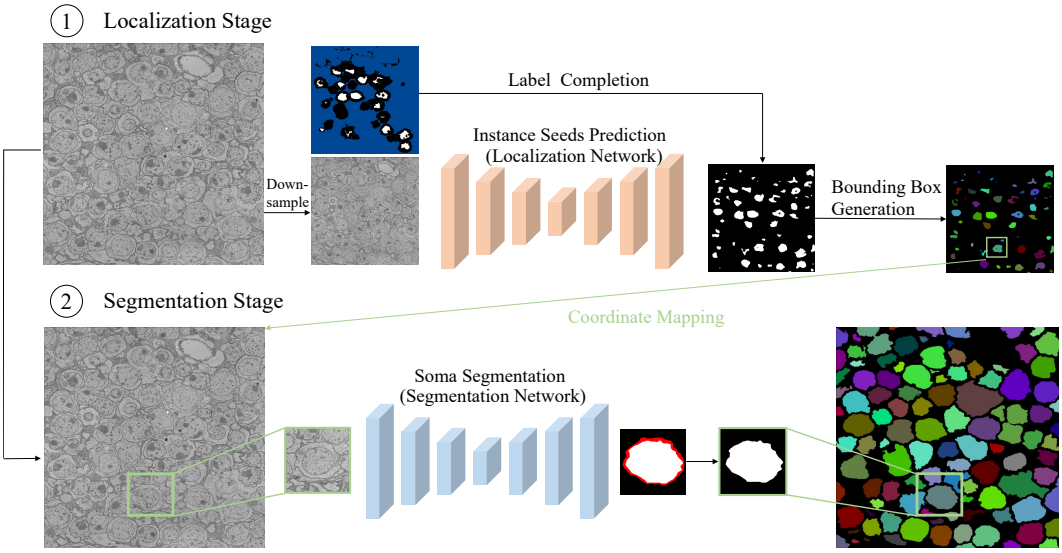

Figure 2: Illustration of our proposed soma instance segmentation algorithm. The localization stage aims to localize somas by predicting instance seeds and generating bounding boxes for them in a given EM block, and the segmentation stage aims to segment somas from the predicted bounding boxes.

4 Two-Stage Soma Instance Segmentation Algorithm

We propose an accurate and efficient soma instance segmentation algorithm. Our algorithm contains two sequential stages, the localization stage and the segmentation stage. The localization stage aims to localize somas by predicting instance seeds and generating bounding boxes for them in a given EM block, and the segmentation stage aims to segment the complete soma (soma-of-interest) located in the center of the predicted bounding boxes. We illustrate the proposed algorithm in Figure 2.

Localization Stage. The localization stage consists of three steps: label completion, bounding box prediction and coordinate mapping. First, we complete the labels for the unannotated areas in our localization label set automatically with a localization network using the U-Net model. Then, we adopt the completed label set to continually train this U-Net model to predict instance seeds of the somas in unseen blocks, and generate the bounding boxes for the somas by the instance seeds. Finally, we map the bounding boxes back to the original blocks. The following segmentation stage segments somas from these bounding boxes. We illustrate the detailed structures of our networks and training details in Appendix A.3 and A.4, respectively.

Label Completion and Instance Seed Prediction. The localization label set is stored after downsampling so that we can feed a full EM block into our localization network within the capacity of GPU memories. The factors of downsampling are 4, 4 and 6 for x-, y- and z-axis, respectively, resulting in a uniform block size of $459 \times 459 \times 62$. Despite that we have labels for a part of the areas in our localization label set, there are still unannotated areas without labels. We thus use our localization label set to train a network to classify the instance seeds of somas and the non-seed areas on the labeled areas, and employ self-training to generate labels for the unannotated areas. First, we ignore the unannotated areas and use the original labels to optimize the network. The optimization only computes loss on the labeled areas. After the optimization tends to converge, we employ the network to predict results for the unannotated areas. We combine both the generated labels and the original labels together to form the final labels for each block. We then adopt the final labels to further optimize the network until the optimization converges.

Bounding Box Generation. After label completion, we obtain a trained classification network. We then use it to predict instance seeds of somas for unseen EM blocks. According to the sizes of the predicted instance seeds, our method automatically generates the 3D bounding boxes for somas that the instance seeds indicate. We set the centroid of the predicted instance seed as the center for

the bounding box. The length of each border of the 3D bounding box is predefined as more than two times the instance seed, which can generally cover the corresponding soma completely. Each bounding box thus indicates a patch that contains a complete soma.

Coordinate Mapping. After obtaining the bounding boxes for somas in a given block, we can localize the somas. However, since the block we use to localize somas has been downsampled before, we have to map the coordinate back to the original block. We multiply the coordinate by the downsampling factors 4, 4 and 6, and adopt the bounding boxes with the multiplied coordinates to localize the somas in the original block.

Segmentation Stage. In the segmentation stage, we train a segmentation network using the U-Net backbone (which can be readily upgraded to more advanced structures such as transformer, as shown in Sec. 6) on our segmentation label set to classify the soma-of-interest, the soma boundary and the rest area from a given patch. Note that the network only predicts the complete soma at the center of the patch, and the rest area is regarded as the background. Based on this special design, we can directly train a 3D model on the partially annotated dataset. Moreover, this design is beneficial for dealing with densely distributed somas, since it is difficult to directly predict the background between adjacent somas, even on fully annotated datasets. During inference, we regard the bounding boxes we obtain in the localization stage as the patches, and employ the trained segmentation network to predict voxel-wise classes for these patches. Since the size of each patch is variable, we set the batch size as 1 during network training. Finally, we take the predicted soma-of-interest class as our soma segmentation results.

5 Parallelized Large-scale Data Processing Pipeline

To cope with the huge amount of EM data of a full adult fly brain and accelerate the computation, we deploy a parallelized, high-throughput data processing pipeline on distributed clusters of CPUs and GPUs, which is shown in Appendix A.5. Overall, our pipeline follows the design of a mainstream distributed processing algorithm (Wu et al., 2021) but is customized based on our local infrastructure and our soma segmentation task. Firstly, we divide the whole FAFB volume into a number of 3D blocks. Secondly, we execute a segmentation procedure to extract somas within each block in parallel. Thirdly, we stitch all block-wise segmentation results to obtain the final reconstruction result for the whole 3D volume.

3D Block Division. Limited by the RAM size, it is impossible to process the whole FAFB volume directly in a single cluster node. We divide the whole 3D volume into overlapping 3D blocks and save them in our purpose-built data center. Then the cluster nodes process these blocks in parallel. Metadata, containing the relative location in the whole 3D volume, accompanies each block. In total, we generate $40,590$ blocks with a size of $1,836 \times 1,836 \times 186$ voxel from the whole FAFB volume. Neighboring blocks share an overlapping region with 212×30 voxels in lateral and axial directions, respectively, which are used for stitching the blocks later.

Intra-block Segmentation. We package these divided blocks into groups, each of which can be distributed to and processed by a computing task with one TITAN XP GPU and corresponding CPUs, RAM, *etc*. We develop a task management front-end to produce and submit these tasks to a task queuing system. All tasks are executed independently to process groups of blocks by the above deep learning-based soma segmentation algorithm. This processing stage requires little inter-process communication. Finally, each task generates a group of intra-block segmentation results and writes them into our data center for the following processing stage.

Inter-block Stitching. To obtain the complete segmentation result of one soma instance across different neighboring blocks, we adopt a hierarchical block stitching algorithm with three steps to stitch the segmentation results efficiently. Firstly, we compute the overlap ratio of each soma between two adjacent blocks. If the overlap ratio exceeds an empirically predetermined threshold, we consider it to be a soma spanning two blocks. In other words, these two somas with different IDs should be merged, *i.e.*, they should be with the same ID. Note that, in order to realize the stitching process in parallel, we store the IDs that need to be merged into a shared list instead of immediately changing the IDs in the current block. Secondly, we unify the pairing of all IDs in the shared list to

ensure that each ID that needs to be merged corresponds to a unique target ID. Finally, we remap the IDs of somas in all blocks according to the shared list. The inter-block stitching step is implemented in a parallel manner, and the details are shown in Appendix A.6.

6 EXPERIMENTAL RESULTS

Visualization Results. We first provide some visualization results for the full brain soma reconstruction, as shown in Appendix A.2. Following the z-axis order in the FAFB dataset, we select three 3D sections of the full brain soma reconstruction results as example, and show the details in two representative regions, respectively. Please refer to the 4K video in Appendix A.1 for the complete reconstruction results.

Quantitative Results. To validate the effectiveness and efficiency of the proposed method, we compare our method with two existing methods for nucleus reconstruction in EM images, *i.e.*, Mu et al. (Mu et al., 2021) and Lin et al. (Lin et al., 2021). To adapt these two methods to our EMADS dataset, we only use the annotated areas to optimize the networks and ignore the unannotated areas. For the second method, we have to downsample the dataset to $1/96$ of the previous volume so that each block can be fed to a standard 3D U-Net as the method requires. Then, the segmentation result is upsampled to the previous scale. In addition, we compare our method with two self-designed baseline methods. Baseline 1 adopts dense sliding bounding boxes in a uniform size on the EM block instead of localization, and utilizes a two-class network to segment the somas and the rest areas. Baseline 2 is similar to Baseline 1, and the only difference is that it adopts a three-class network for segmentation. After the dense sliding finishes, the segmented somas are endowed with instances by their connected components in the block.

Evaluation Datasets and Metrics. We fully annotate the somas in three additional EM blocks named A, B and C, each of which contains about 50 somas. They are non-overlapping with the 10 training blocks in EMADS, serving as the evaluation blocks. The somas in different blocks have different sizes and morphologies. The background areas are also of different types. We adopt A and B as our test blocks, and C as the validation block for the network training. We adopt 3D Average Precision metrics (Wei et al., 2020), *i.e.*, mAP and mAP_{50} as the instance segmentation metrics to evaluate the instance localization and segmentation performance. In addition, we also use the Jaccard score (Niwattanakul et al., 2013) as the semantic segmentation metric to evaluate the semantic segmentation performance. Since some somas located at the boundary of the blocks are hard to annotate, we exclude them when we calculate segmentation metrics.

Comparison with Existing Methods. As shown in Table 1, compared with the two existing methods for nucleus reconstruction, our method with the U-Net backbone achieves much better performance in terms of the instance segmentation metrics mAP and mAP_{50},due to its special design tailored for the soma reconstruction task. It should be noticed that, however, these two existing methods perform not bad in terms of the semantic segmentation metric, which suggests that they mainly fail to distinguish between soma instances. This can be verified by visual comparison examples in Figure 3. As can be seen, our results with fewer segmentation errors provide more accurate distribution and morphology information of somas.

Comparison with Baseline Methods. Compared with Baseline 1 and Baseline 2, our method is about 6 times faster and achieves better performance, which validates the efficiency of the localization stage in our method. We also provide visual comparison examples with the two baseline methods in Appendix A.7. To further demonstrate the generality of our method, we upgrade the 3D U-Net used in the segmentation stage to the state-of-the-art backbone based on the transformer architecture, *i.e.*, Swin-UNETER (Tang et al., 2022). The result of 'Ours-Swin' proves that, thanks to the flexibility of our two-stage segmentation algorithm, a more powerful backbone network brings new performance improvements.

Ablation on seed size. When we make our localization label set, we need to erode the annotated somas to obtain the instance seeds. The sizes of the instance seeds are different when we adopt different erosion operations. We thus conduct an ablation study on seed size. The seed size denotes the maximum number of voxels a seed contains. The ablation results are shown in Table 2, which demonstrates that the seed size of 50 provides the best performance.

Table 1: Quantitative comparison results with existing and baseline methods.

Method	Test block A			Test block B			Average			
	mAP	mAP_{50}	Jacc.	mAP	mAP_{50}	Jacc.	mAP	mAP_{50}	Jacc.	time
Mu et al.	0.045	0.212	0.579	0.072	0.219	0.578	0.059	0.216	0.579	63s
Lin et al.	0.017	0.096	0.420	0.020	0.093	0.397	0.019	0.095	0.409	32s
Baseline 1	0.213	0.699	0.587	0.179	0.680	0.524	0.196	0.690	0.556	960s
Baseline 2	0.226	0.695	0.592	0.242	0.709	0.558	0.234	0.702	0.575	1142s
Ours-UNET	0.301	0.713	0.638	0.302	**0.721**	**0.590**	0.301	0.717	**0.614**	178s
Ours-Swin	**0.420**	**0.853**	**0.650**	**0.303**	0.614	0.474	**0.362**	**0.734**	0.562	158s

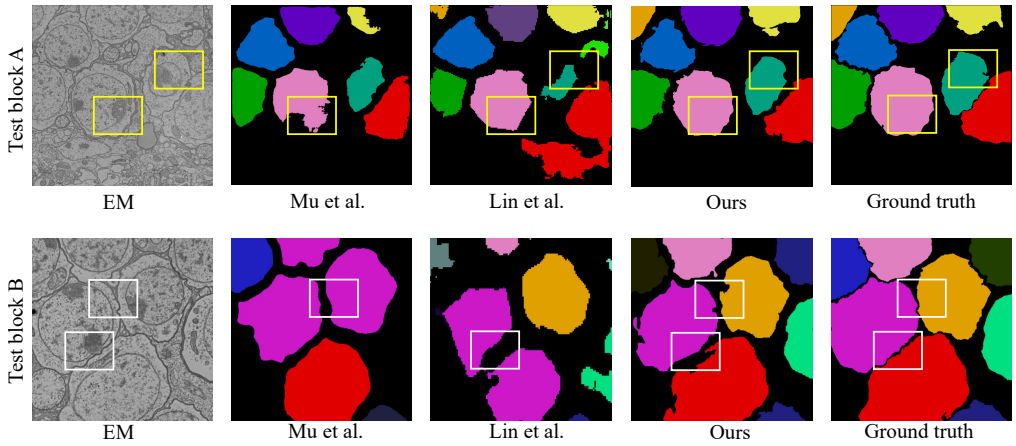

Figure 3: Visual comparison results of our method and two existing methods. The yellow box and the white box indicate the segmentation and localization errors.

7 STATISTICS

We compile statistics for the reconstructed somas in the full adult fly brain in four aspects: count, size, distribution and morphology. To make the reconstructed somas correspond to the neurobiological structures in the brain, we partition the full brain into four types of cubic regions: the optic lobes, and the central brain A, B and C as shown in the top left in Figure 4. We compute the statistical results for each of them. More details of the brain partitions are shown in Appendix A.8. According to the official nomenclature of brain structures for the drosophila (Ito et al., 2014), the central brain A contains superior neuropils (SNP), mushroom bodies (MB), inferior neuropils (INP), central complex (CX) and lateral horn (LH); the central brain B contains ventrolateral neuropils (VLNP), ventromedial neuropils (VMNP), antennal lobes (AL), lateral complex (LX); the central brain C contains periesophageal neuropils (PENP), gnathal ganglia (GNG) and external nerves.

Soma Count. The number of soma instances we reconstruct from the full adult fly brain is $116,761$. This number is basically consistent with the result in the recent nucleus reconstruction work (Mu et al., 2021) and the result reported by a traditional method Godfrey et al. (2021). Regarding the count in each region, the optic lobes contain the majority of somas in the full brain with a number of $76,316$. The central brain A, B and C contain $18,313$, $14,424$ and $7,708$ somas, respectively. The number of somas in the central brain A is obviously more than that of B and C due to the abundant Kenyon cells in the mushroom body (Ito et al., 2014).

Soma Size. We illustrate the counts of somas in different sizes of each region in the second row of Figure 4. The largest soma in the full brain has a size of $980\ \mu m^3$, and most of the soma sizes are

Table 2: Ablation on seed size.

Size	Test block A			Test block B			Average		
	mAP	mAP_{50}	Jacc.	mAP	mAP_{50}	Jacc.	mAP	mAP_{50}	Jacc.
80	0.261	0.696	0.606	0.276	0.728	0.574	0.269	0.712	0.590
50 (Ours)	**0.301**	0.713	**0.638**	**0.302**	**0.721**	**0.590**	**0.302**	**0.717**	**0.614**
10	0.284	**0.715**	0.625	0.246	0.667	0.579	0.265	0.691	0.602

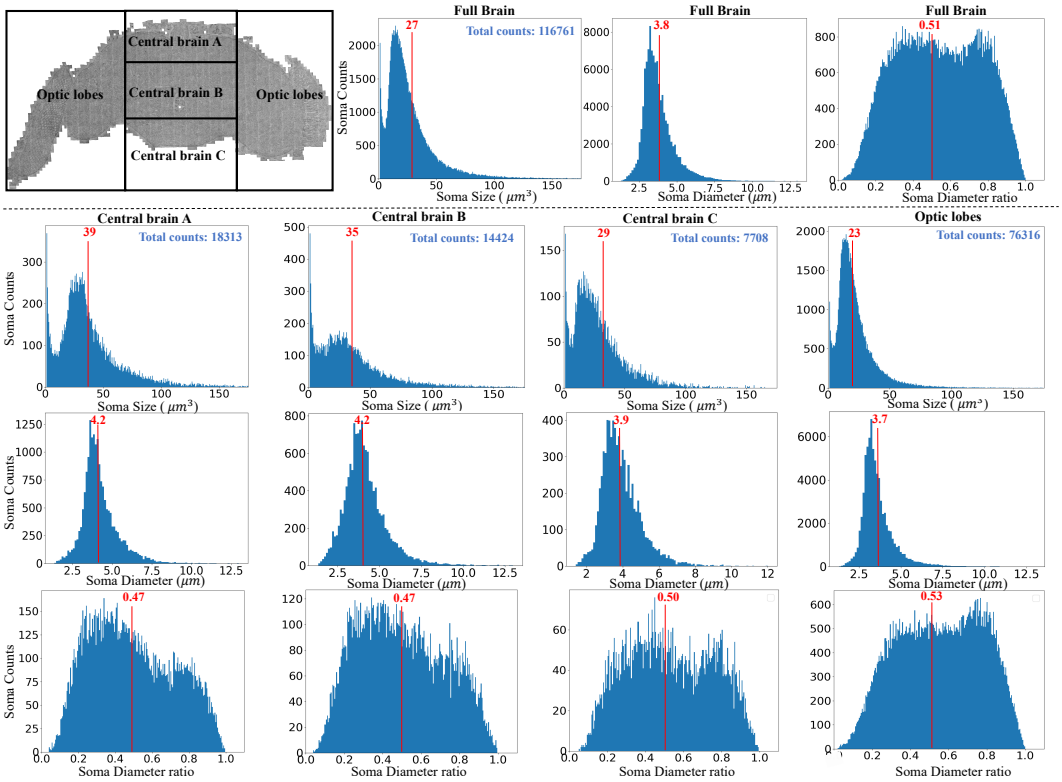

Figure 4: Statistics of soma sizes and diameters in different regions of the full brain. The first row illustrates our separated four types of regions and the statistical results on the full brain. The second row illustrates the soma count in different sizes, the third row illustrates the soma count in different soma diameters, and the fourth row illustrates the soma count in different diameter ratios. The red line highlights the mean value of each statistic.

less than 75 μm^3. The mean soma size in the full brain is 27 μm^3. The mean sizes of somas in the central brain A, B, C and the optic lobes are 39, 35, 29 and 23 μm^3, respectively.

Soma Distribution. From Appendix A.2, we can observe that the somas mainly locate at the rind of the full brain, and a few somas locate at the center of the brain. The results in Figure 4 demonstrate that the optic lobes have the most somas, while the central brain C has the fewest somas. The statistical results are consistent with the visualization results.

Soma Morphology. Somas in different morphologies have different diameters (*i.e.*, the maximum length of its spatial size). We count the somas in different diameters. The soma diameters are basically consistent in the four regions, as shown in the third row in Figure 4. The somas in the optic lobes have the smallest mean diameter, which corresponds to the fact that there are much more neural stem cells that are undifferentiated in this region than in the central brain A, B and C, as introduced in (Ito et al., 2013). Moreover, we count somas in different diameter ratios (*i.e.*, ratios between the maximum and the minimum lengths of its spatial size), which reflects the roundness of the sphere. As shown in the fourth row in Figure 4, the somas in the optic lobes have the maximum mean diameter ratio and the morphology of these somas is more similar to the sphere.

8 Conclusion

In this paper, we have made one of the first efforts to develop an accurate and efficient soma reconstruction method for a full adult fly brain. We first contribute the EMADS dataset with fine-grained annotation on somas, based on which we then propose an efficient, two-stage deep learning algorithm for accurate soma instance segmentation and implement it on the full brain by deploying a parallelized large-scale data processing pipeline. We hope this work could benefit the future study of the drosophila neural system.

REFERENCES

Birgitte Bo Andersen, Hans Jørgen G Gundersen, and Bente Pakkenberg. Aging of the human cerebellum: a stereological study. *Journal of Comparative Neurology*, 466(3):356–365, 2003.

J Douglas Armstrong and Jano I van Hemert. Towards a virtual fly brain. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1896):2387–2397, 2009.

J Alexander Bae, Mahaly Baptiste, Agnes L Bodor, Derrick Brittain, JoAnn Buchanan, Daniel J Bumbarger, Manuel A Castro, Brendan Celii, Erick Cobos, Forrest Collman, et al. Functional connectomics spanning multiple areas of mouse visual cortex. *bioRxiv*, 2021.

Daniel R Berger, H Sebastian Seung, and Jeff W Lichtman. Vast (volume annotation and segmentation tool): efficient manual and semi-automatic labeling of large 3d image stacks. *Frontiers in neural circuits*, 12:88, 2018.

Rolf Bodmer and Yuh Nung Jan. Morphological differentiation of the embryonic peripheral neurons in drosophila. *Roux's archives of developmental biology*, 196(2):69–77, 1987.

Thomas Burne, E Scott, Bruno van Swinderen, M Hilliard, Judith Reinhard, Charles Claudianos, D Eyles, and John McGrath. Big ideas for small brains: what can psychiatry learn from worms, flies, bees and fish? *Molecular psychiatry*, 16(1):7–16, 2011.

Sebastian Busch, Mareike Selcho, Kei Ito, and Hiromu Tanimoto. A map of octopaminergic neurons in the drosophila brain. *Journal of Comparative Neurology*, 513(6):643–667, 2009.

Sebastian Cachero, Aaron D Ostrovsky, Y Yu Jai, Barry J Dickson, and Gregory SXE Jefferis. Sexual dimorphism in the fly brain. *Current biology*, 20(18):1589–1601, 2010.

Ann-Shyn Chiang, Chih-Yung Lin, Chao-Chun Chuang, Hsiu-Ming Chang, Chang-Huain Hsieh, Chang-Wei Yeh, Chi-Tin Shih, Jian-Jheng Wu, Guo-Tzau Wang, Yung-Chang Chen, et al. Three-dimensional reconstruction of brain-wide wiring networks in drosophila at single-cell resolution. *Current biology*, 21(1):1–11, 2011.

Michael H Dickinson and Florian T Muijres. The aerodynamics and control of free flight manoeuvres in drosophila. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371 (1704):20150388, 2016.

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.

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.

Miguel Á García-Cabezas, Yohan J John, Helen Barbas, and Basilis Zikopoulos. Distinction of neurons, glia and endothelial cells in the cerebral cortex: an algorithm based on cytological features. *Frontiers in neuroanatomy*, 10:107, 2016.

George L Gerstein. Analysis of firing patterns in single neurons. *Science*, 131(3416):1811–1812, 1960.

Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp. 315–323. JMLR Workshop and Conference Proceedings, 2011.

R Keating Godfrey, Mira Swartzlander, and Wulfila Gronenberg. Allometric analysis of brain cell number in hymenoptera suggests ant brains diverge from general trends. *Proceedings of the Royal Society B*, 288(1947):20210199, 2021.

Eyal Gruntman and Glenn C Turner. Integration of the olfactory code across dendritic claws of single mushroom body neurons. *Nature neuroscience*, 16(12):1821–1829, 2013.

Volker Hartenstein. Morphological diversity and development of glia in drosophila. *Glia*, 59(9): 1237–1252, 2011.

Suzana Herculano-Houzel and Roberto Lent. Isotropic fractionator: a simple, rapid method for the quantification of total cell and neuron numbers in the brain. *Journal of Neuroscience*, 25(10): 2518–2521, 2005.

Kei Ito, Kazunori Shinomiya, Masayoshi Ito, J Douglas Armstrong, George Boyan, Volker Hartenstein, Steffen Harzsch, Martin Heisenberg, Uwe Homberg, Arnim Jenett, et al. A systematic nomenclature for the insect brain. *Neuron*, 81(4):755–765, 2014.

Masayoshi Ito, Naoki Masuda, Kazunori Shinomiya, Keita Endo, and Kei Ito. Systematic analysis of neural projections reveals clonal composition of the drosophila brain. *Current Biology*, 23(8): 644–655, 2013.

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):605–610, 2018.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Roberto Lent, Frederico AC Azevedo, Carlos H Andrade-Moraes, and Ana VO Pinto. How many neurons do you have? some dogmas of quantitative neuroscience under revision. *European Journal of Neuroscience*, 35(1):1–9, 2012.

Peter H Li, Larry F Lindsey, Michał Januszewski, Mike Tyka, Jeremy Maitin-Shepard, Tim Blakely, and Viren Jain. Automated reconstruction of a serial-section em drosophila brain with flood-filling networks and local realignment. *Microscopy and Microanalysis*, 25(S2):1364–1365, 2019.

Zudi Lin, Donglai Wei, Mariela D Petkova, Yuelong Wu, Zergham Ahmed, Silin Zou, Nils Wendt, Jonathan Boulanger-Weill, Xueying Wang, Nagaraju Dhanyasi, et al. Nucmm dataset: 3d neuronal nuclei instance segmentation at sub-cubic millimeter scale. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 164–174. Springer, 2021.

Sergio Luengo-Sanchez, Concha Bielza, Ruth Benavides-Piccione, Isabel Fernaud-Espinosa, Javier DeFelipe, and Pedro Larrañaga. A univocal definition of the neuronal soma morphology using gaussian mixture models. *Frontiers in Neuroanatomy*, 9:137, 2015.

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*, 366(6469):eaay3134, 2019.

Shang Mu, Szi-chieh Yu, Nicholas L Turner, Claire E McKellar, Sven Dorkenwald, Forrest Collman, Selden Koolman, Merlin Moore, Sarah Morejohn, Ben Silverman, et al. 3d reconstruction of cell nuclei in a full drosophila brain. *bioRxiv*, 2021.

Suphakit Niwattanakul, Jatsada Singthongchai, Ekkachai Naenudorn, and Supachanun Wanapu. Using of jaccard coefficient for keywords similarity. In *Proceedings of the international multiconference of engineers and computer scientists*, volume 1, pp. 380–384, 2013.

Tyler A Ofstad, Charles S Zuker, and Michael B Reiser. Visual place learning in drosophila melanogaster. *Nature*, 474(7350):204–207, 2011.

Tomoko Ohyama, Casey M Schneider-Mizell, Richard D Fetter, Javier Valdes Aleman, Romain Franconville, Marta Rivera-Alba, Brett D Mensh, Kristin M Branson, Julie H Simpson, James W Truman, et al. A multilevel multimodal circuit enhances action selection in drosophila. *Nature*, 520(7549):633–639, 2015.

David Owald and Scott Waddell. Olfactory learning skews mushroom body output pathways to steer behavioral choice in drosophila. *Current opinion in neurobiology*, 35:178–184, 2015.

Hania J Pavlou and Stephen F Goodwin. Courtship behavior in drosophila melanogaster: towards a 'courtship connectome'. *Current Opinion in Neurobiology*, 23(1):76–83, 2013.

Karlheinz Rein, Malte Zöckler, Michael T Mader, Cornelia Grübel, and Martin Heisenberg. The drosophila standard brain. *Current Biology*, 12(3):227–231, 2002.

Alexander Shapson-Coe, Michał Januszewski, Daniel R Berger, Art Pope, Yuelong Wu, Tim Blakely, Richard L Schalek, Peter H Li, Shuohong Wang, Jeremy Maitin-Shepard, et al. A connectomic study of a petascale fragment of human cerebral cortex. *BioRxiv*, 2021.

Julie H Simpson. Mapping and manipulating neural circuits in the fly brain. *Advances in genetics*, 65:79–143, 2009.

Shin-ya Takemura, C Shan Xu, Zhiyuan Lu, Patricia K Rivlin, Toufiq Parag, Donald J Olbris, Stephen Plaza, Ting Zhao, William T Katz, Lowell Umayam, et al. Synaptic circuits and their variations within different columns in the visual system of drosophila. *Proceedings of the National Academy of Sciences*, 112(44):13711–13716, 2015.

Shin-ya Takemura, Yoshinori Aso, Toshihide Hige, Allan Wong, Zhiyuan Lu, C Shan Xu, Patricia K Rivlin, Harald Hess, Ting Zhao, Toufiq Parag, et al. A connectome of a learning and memory center in the adult drosophila brain. *Elife*, 6:e26975, 2017.

Yucheng Tang, Dong Yang, Wenqi Li, Holger R Roth, Bennett Landman, Daguang Xu, Vishwesh Nath, and Ali Hatamizadeh. Self-supervised pre-training of swin transformers for 3d medical image analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20730–20740, 2022.

Wynand van der Goes van Naters and John R Carlson. Receptors and neurons for fly odors in drosophila. *Current biology*, 17(7):606–612, 2007.

Donglai Wei, Zudi Lin, Daniel Franco-Barranco, Nils Wendt, Xingyu Liu, Wenjie Yin, Xin Huang, Aarush Gupta, Won-Dong Jang, Xueying Wang, et al. Mitoem dataset: large-scale 3d mitochondria instance segmentation from em images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 66–76. Springer, 2020.

Jingpeng Wu, William M Silversmith, Kisuk Lee, and H Sebastian Seung. Chunkflow: hybrid cloud processing of large 3d images by convolutional nets. *Nature Methods*, 18(4):328–330, 2021.

Zhihao Zheng, J Scott Lauritzen, Eric Perlman, Camenzind G Robinson, Matthew Nichols, Daniel Milkie, Omar Torrens, John Price, Corey B Fisher, Nadiya Sharifi, et al. A complete electron microscopy volume of the brain of adult drosophila melanogaster. *Cell*, 174(3):730–743, 2018.

A APPENDIX

Here we provide additional visual results and implementation details that have not been presented in the main paper.

A.1 VIDEO VISUALIZATION

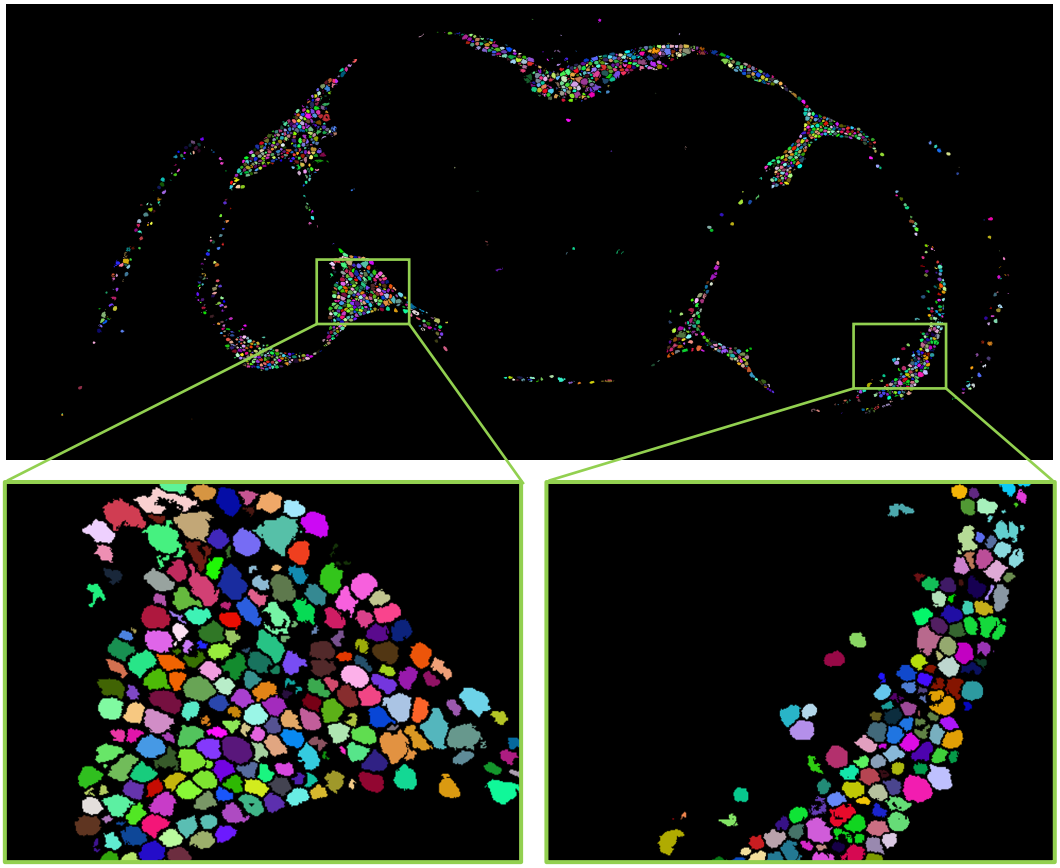

Figure 5: An image example of our provided image sequence with two areas zoomed in.

We provide a 2D image sequence to overall visualize the soma reconstruction results in a full adult fruit fly brain. Specifically, we downsample the full brain reconstruction result by the factors of 10, 10, 4 at the x-axis and y-axis, and visualize the result along the z-axis. Since our reconstruction is based on the physical resolution 16nm $\times$ 16nm $\times$ 40nm ($x \times y \times z$), the physical resolution of this overall full brain visualization is 160nm $\times$ 160nm $\times$ 160nm ($x \times y \times z$). An image example with two areas zoomed in is shown as Figure 5.

With this image sequence, we can obtain a panoramic view for the full brain reconstruction result. In addition, we also make a 4K video based on this image sequence. The 4K video link is: https://drive.google.com/drive/folders/1KT3f2gVVcGtXjklA-E7G31kkfS4rZKM3. We set the time duration of each image as 0.01 second, and stack the image sequence into a video chronological. Each color denotes a soma instance. For convenience, we also provide a light video demo $full_brain_soma_demo.mp4$ in our supplementary material.

A.2 3D VISUALIZATION

We visualize the full brain soma reconstruction result of three 3D sections in Figure 6. Also, we further zoom in a section with a higher factor to see the reconstructed somas in detail in Figure 7.

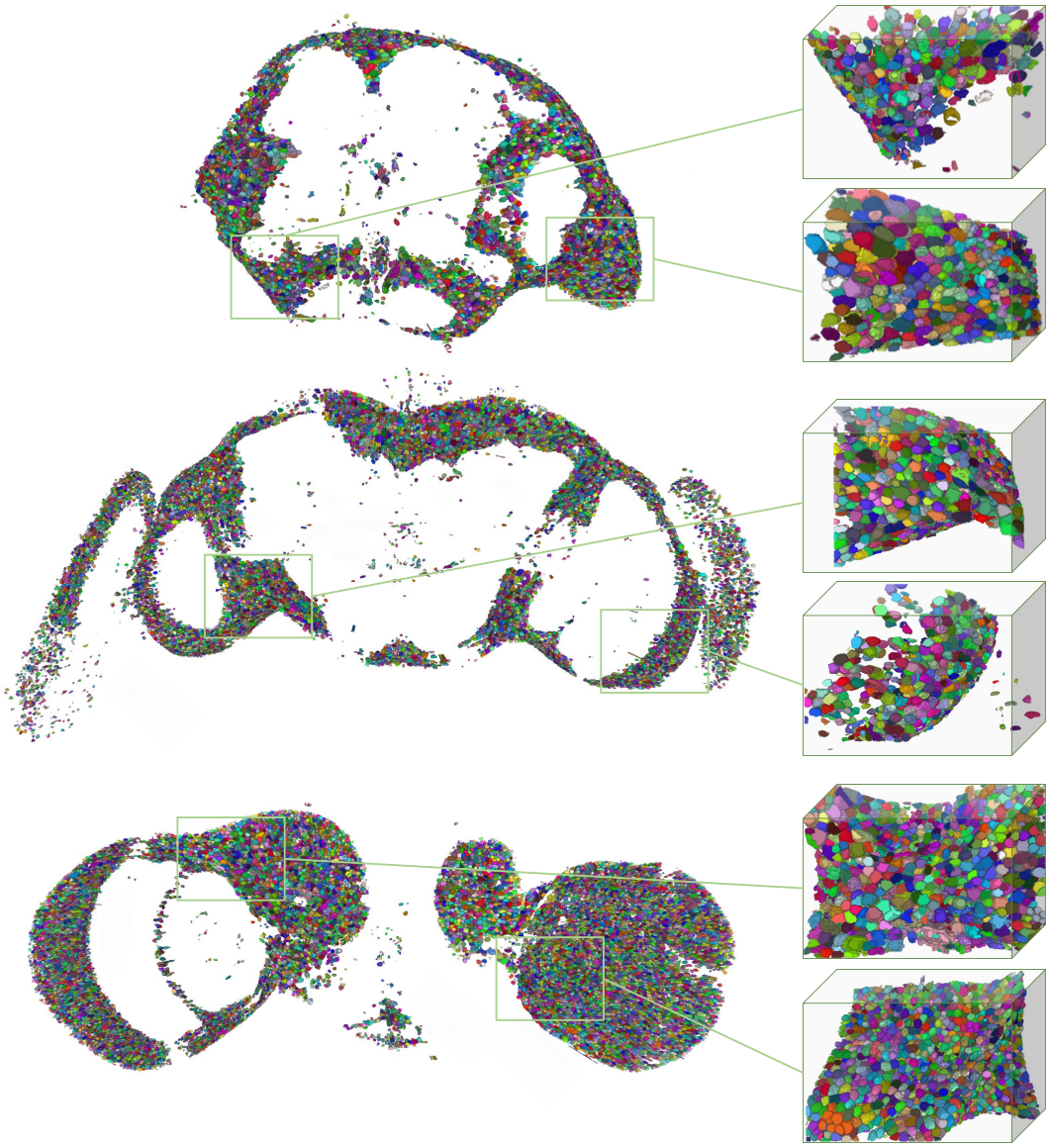

Figure 6: Visualization of our soma reconstruction for a full adult drosophila brain in three 3D sections. Each color instance corresponds to a reconstructed soma.

A.3 Network Structures

We illustrate the detailed structures of our deep networks in Figure 8 which are introduced in Section 3 in our main paper. Both the localization and segmentation networks are based on 3D U-Net, and they have the same structure except for the last convolution layer. Moreover, each convolution layer except for the last one is followed by a ReLU layer (Glorot et al., 2011).

A.4 Training Details

We train the localization network using a learning rate of 5×10^{-4} and a batch size of 4 on four NVIDIA TITAN XP GPU for 500 epochs, and train the segmentation network using a learning rate of 1×10^{-4} and a batch size of 1 on one NVIDIA TITAN XP GPU for 200 epochs. We optimize both of the networks by the Adam optimizer (Kingma & Ba, 2014) with $\beta_1 = 0.9$ and $\beta_2 = 0.99$.

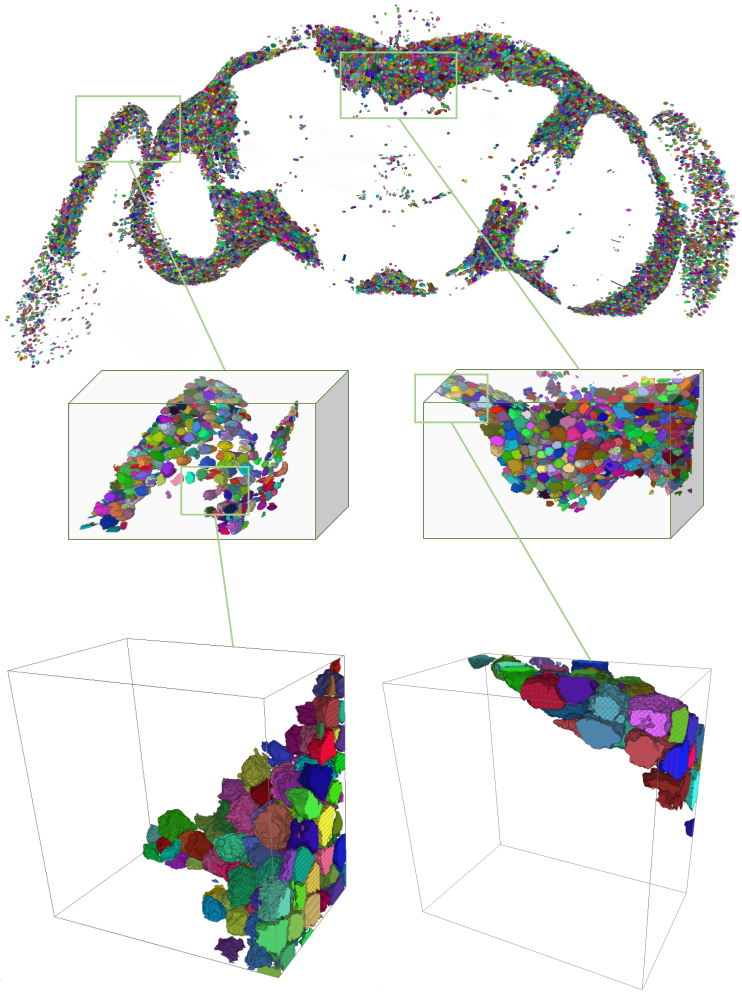

Figure 7: A detailed visualization example with two times zooming in.

A.5 Parallelized Large-scale Data Processing Pipeline

The parallelized large-scale data processing pipeline is shown in Figure 9.

A.6 Details of Parallel Inter-block Stitching

This section elaborates our hierarchical block stitching algorithm, which is used for parallel inter-block stitching in the large-scale data processing pipeline. As shown in Figure 10, we divide the relations of these neighboring 3D blocks in the full brain into three types according to their relative positions. We stitch these blocks in different directions pair by pair. Specifically, for each direction, we divide neighboring blocks into two sequences, the odd and the even, which are indicated by the red and white double-headed arrows, respectively. We first stitch the blocks from the odd sequence, and then we stitch the blocks from the even sequence. The two sequences can be processed independently without interfering with each other to accelarate the pipeline. As an example of the volume consisting of $6 \times 6 \times 6$ blocks in Figure 10, we stitch 180 block pairs in each direction to obtain the final result. With sufficient computing resource (enough CPUs), we can first stitch 90 block pairs in the odd sequence in parallel, and then stitch the other 90 block pairs in the even sequence in parallel.

A.7 Visual comparison examples with the two baseline methods.

Visual comparison examples with the two baseline methods is shown in Figure 11.

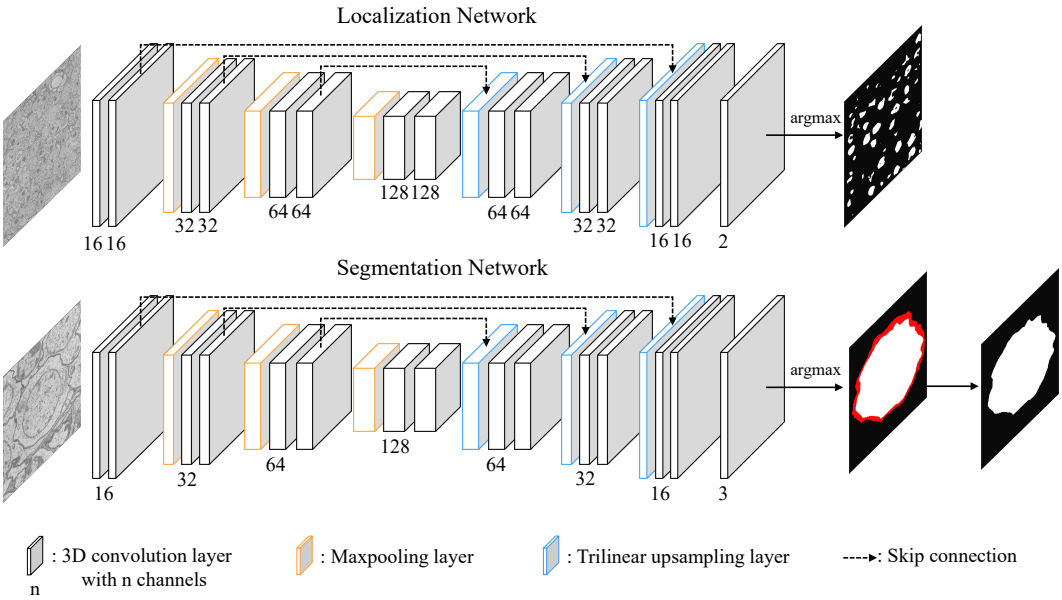

Figure 8: Illustration of the detailed structures of our deep networks. For simplicity and clarity, we illustrate the input and the output image as a slice of the 3D volume.

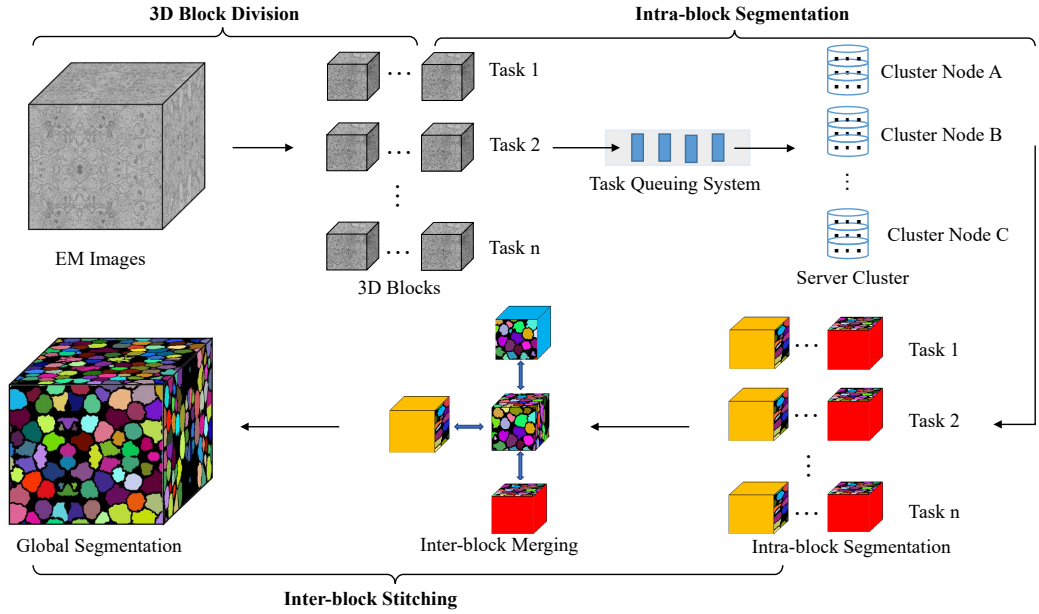

Figure 9: Overview of our parallelized large-scale data processing pipeline.

A.8 Full Brain Partition

As introduced in Section 6 in our main paper, we separate the full brain into four types of cubic regions for statistics. As introduced in Section 2 in our main paper, the full brain data is stored by our defined 3D image blocks, and the image resolution of each block is $1836 \times 1836 \times 186$. The full brain data is stacked by $41 \times 22 \times 45$ blocks ($x \times y \times z$). Our sequence is listed in the $z-$axis order with 45 images, and the size of each image is 41×22. Each pixel in an image denotes a block. The pink, blue and purple color denotes that the block belongs to the central brain A, B and C region, respectively. The green color denotes that the block belongs to the optical lobes.

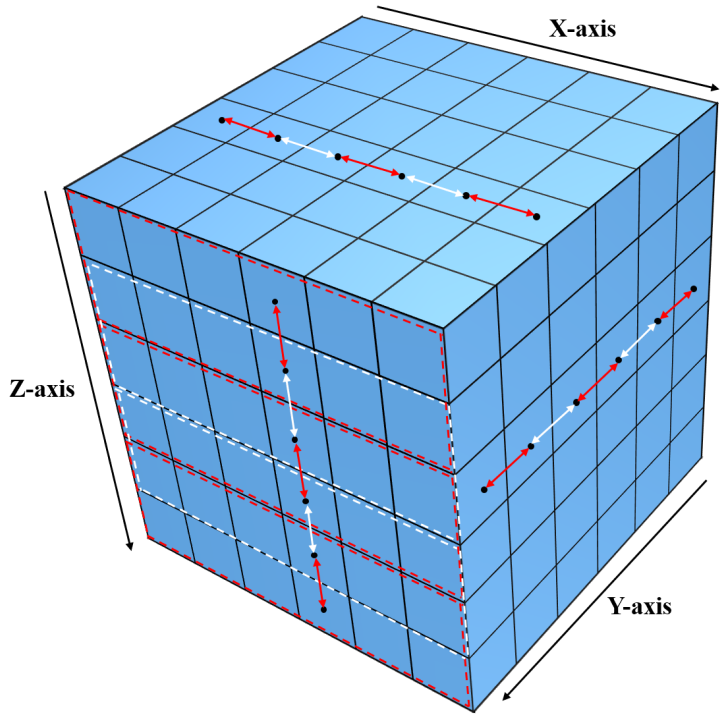

Figure 10: Illustration of our stitching algorithm. The red and white double-headed arrows indicate the pairwise blocks in the odd and even sequences, respectively.

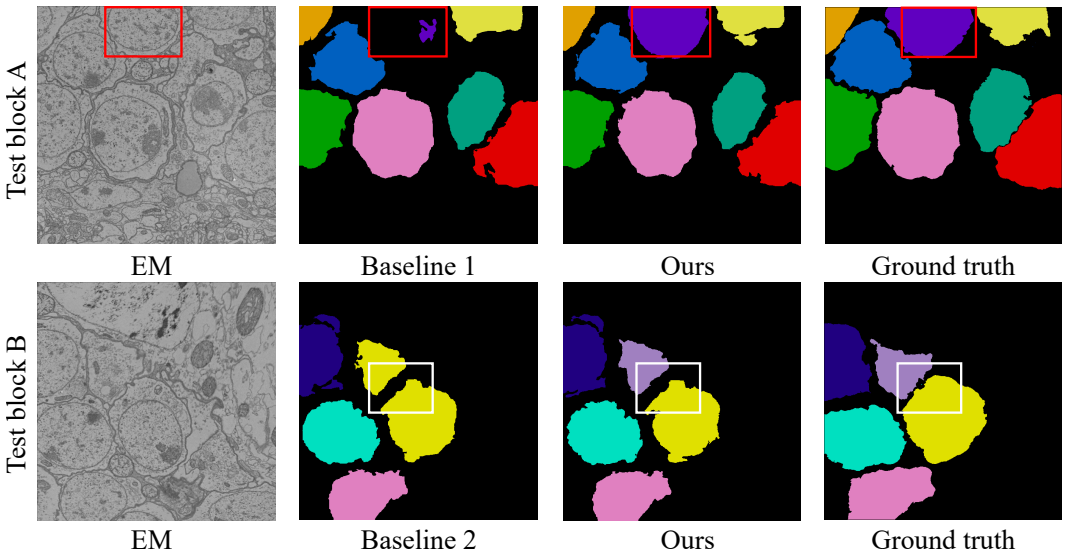

Figure 11: Visual comparison results of our method and two baseline methods.

We illustrate one image of this sequence as an example and the stacked image sequence for brain partition in Figure 12.

A.9 Code and Dataset

Our code link is: https://anonymous.4open.science/r/Accurate-and-Efficient-Soma-Reconstruction-in-a-Full-Adult-Fruit-Fly-Brain-4DD7.

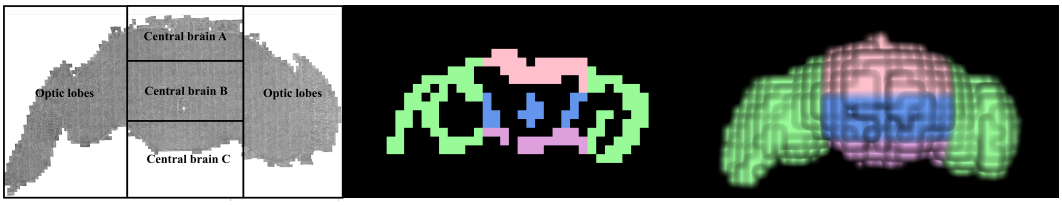

Figure 12: Illustration of our full brain partition. Left: our brain partition; Middle: an image example in the sequence for brain partition ; Right: the stacked 3D image sequence for brain partition.

We provide a test case of our localization label set and segmentation label set separately in the code link with corresponding inference models, and the full codes and the annotated dataset will be organized and released soon.