



Petabyte-Scale Multi-Morphometry of Single Neurons for Whole Brains

Shengdian Jiang^{1,2} · Yimin Wang^{1,3} · Lijuan Liu¹ · Liya Ding¹ · Zongcai Ruan¹ · Hong-Wei Dong⁴ · Giorgio A. Ascoli⁵ · Michael Hawrylycz⁶ · Hongkui Zeng⁶ · Hanchuan Peng¹

Accepted: 21 January 2022 / Published online: 19 February 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Recent advances in brain imaging allow producing large amounts of 3-D volumetric data from which morphometry data is reconstructed and measured. Fine detailed structural morphometry of individual neurons, including somata, dendrites, axons, and synaptic connectivity based on digitally reconstructed neurons, is essential for cataloging neuron types and their connectivity. To produce quality morphometry at large scale, it is highly desirable but extremely challenging to efficiently handle petabyte-scale high-resolution whole brain imaging database. Here, we developed a multi-level method to produce high quality somatic, dendritic, axonal, and potential synaptic morphometry, which was made possible by utilizing necessary petabyte hardware and software platform to optimize both the data and workflow management. Our method also boosts data sharing and remote collaborative validation. We highlight a petabyte application dataset involving 62 whole mouse brains, from which we identified 50,233 somata of individual neurons, profiled the dendrites of 11,322 neurons, reconstructed the full 3-D morphology of 1,050 neurons including their dendrites and full axons, and detected 1.9 million putative synaptic sites derived from axonal boutons. Analysis and simulation of these data indicate the promise of this approach for modern large-scale morphology applications.

Keywords Multi-morphometry · Neuron reconstruction · Whole brain imaging data · Data and workflow management

Introduction

Reconstructing the complete 3-D shape or morphology of a neuron, including its dendrites and axons in their entirety, as well as finer structures such as the somata, dendritic spines, and axonal terminal boutons, is recognized as a crucial step to profile the myriad types of neurons in brains (BICCN, 2021; Winnubst et al., 2019; Gong et al., 2013; Peng et al., 2021). This technique, which we refer to as *Multi-Morphometry*, has begun to generate intriguing information and hypotheses about brain circuits at the single-neuron/single-synapse level (Iascone et al., 2020; Kim et al., 2012; Parekh & Ascoli, 2013).

Mammalian brains, of at least hundreds of cubic millimeters in volume, are very large when sub-micrometer resolution imaging is used to acquire 3-D volumetric image datasets at the whole-brain scale. A fundamental challenge in multi-morphometry is that sub-micrometer resolution is necessary to analyze synaptic patterns in a neuron's arborization, while whole-brain scale is essential to delineate long projecting axonal arbors (Ropireddy et al., 2011). As a result, even for

Shengdian Jiang, Yimin Wang, and Lijuan Liu contributed equally to this study.

✉ Hanchuan Peng
h@braintell.org

¹ SEU-ALLEN Joint Center, Institute for Brain and Intelligence, Southeast University, Nanjing, Jiangsu, China

² School of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu, China

³ School of Computer Engineering and Science, Shanghai University, Shanghai, China

⁴ Department of Neurobiology, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA

⁵ Volgenau School of Engineering, George Mason University, Fairfax, VA 22030, USA

⁶ Allen Institute for Brain Science, Seattle, WA 98109, USA

the mouse brain, a widely used model system of mammalian brains, a typical 3-D brain-image dataset will have tens of teravoxels in volume (Econo et al., 2016; Gong et al., 2016). On the other hand, neurons have a very complicated tree-like shape, and are often labelled and visualized sparsely using chemical (Zeng, 2018; Zingg et al., 2014), transgenic (Rotolo et al., 2008) or viral approaches (Aransay et al., 2015; Karube et al., 2004). The number of morphologically distinguishable neurons per brain is often limited. Therefore, to understand the vast complexity and variation of neurons, it is crucial to obtain a large collection of brain image datasets (Micheva & Smith, 2007; Sivagnanam et al., 2015). As each voxel is often stored as one or more bytes, the multi-morphometry problem arises as a petabyte-computing challenge, and as a paramount task for current bioimage informatics applications and technologies (Eliceiri et al., 2012; Meijering et al., 2016; Myers, 2012; Swedlow et al., 2003).

There is a long history of reconstructing individual neurons' morphology with image analysis (Acciai et al., 2016; Cohen et al., 1994). Subneuronal structures including somata, spines and boutons have also been segmented and analyzed from images (Iascone et al., 2020; Cheng et al., 2019; Bass et al., 2017; Gala et al., 2017; Yan et al., 2013; Peng et al., 2010; Liu et al., 2019). This is a challenge of high community interest. A number of algorithms have been examined and compared against each other in public initiatives, e.g. DIADEM (Gillette et al., 2011) or in the global collaborative BigNeuron initiative (Peng et al., 2015). However, most existing methods are applicable only to smaller image datasets and partial neuronal structures. For individual mammalian brain datasets, technologies that can handle teravoxels of image volume to trace millimeters long neurite fibers emerged only recently, including TeraFly (Bria et al., 2016), UltraTracer (Peng et al., 2017), BigData-Viewer (Pietzsch et al., 2015), TeraVR (Wang et al., 2019), and GTree (Zhou et al., 2021). Manual and semi-automatic methods were used to trace neurons' full skeletons in Janelia's MouseLight project (Winnubst et al., 2019). Yet, it is largely an open problem how to scale up all these approaches to handle petabyte-scale multi-morphometry challenge that is becoming a compelling reality as whole-brain screening projects involving increasingly larger and more complicated animal models are being carried out internationally (BRAIN Initiative (<https://braininitiative.nih.gov/>), Blue Brain Project (Markram, 2006), etc.).

Methods

Hardware and Software Platform

Our system is designed for the task of multi-morphometry data generation from petabyte-scale whole-brain imaging

database. Built upon customized hardware infrastructure and software tools, the system is capable of organizing whole-brain imaging datasets, generating multi-morphometry data, managing data and workflow, and visualization and analysis of the generated data (Supplementary Fig. 1).

The hardware infrastructure includes the following parts. VR-equipped annotation workstations are used for data visualization, interactive neuron reconstruction, proofreading, etc. A petabyte-scale storage is configured to store the whole-brain imaging datasets, while the multi-morphometry data is managed using cloud-based storage. A computing cluster is deployed for parallel execution of batch work assignments. Moreover, a wall-mounted display array is also available for monitoring the data generation status. The storage server and the computing cluster are connected with a 100 Gbps wired local area network for peta-scale data storage and parallel computing. The annotation workstations, cloud storage, and monitor system reside in a 10 Gbps local area network.

There are four major software components in the *MorphoHub* software package. The *MorphoHub*-DBMS is the core of the entire system which is responsible for the coordination of the overall data generation process. The DendriteGenerator is in charge for parallel generation of dendritic arbors. The L0Generator is useful for creating compact image representation of the reconstructions. The BoutonGenerator is capable of automatically detection of the synaptic boutons located on axons. These components are developed as plugins of Vaa3D (Peng et al., 2014), thus making *MorphoHub* cross-platform and deployable on various operating systems.

Hardware Configuration

VR-equipped annotation workstations are with Intel Core i7-7700 CPU @ 3.60 GHz, 64 GB memory, NVIDIA GeForce GTX1070 GPU, Windows 10 64-bit edition, and HTC Vive VR headsets. A storage server with 1.6 PB storage capacity and 256 GB memory is used as the Petabyte-scale data storage. A high-performance server with 7 computing nodes and 196 CPU kernels is used as the computing cluster. The wall-mounted display contains a 5*3 array of 28-inch 1920 × 1080 pixel displays and a controller workstation.

MorphoHub-DBMS

The *MorphoHub*-DBMS manages the generation of multi-morphometry data by either collaborating teams or automatic routines. All the morphometry data files are organized in the *MorphoHub*-Database, where they follow consistent naming rule and are assigned with unique IDs. The validity of the database is monitored by an error checking routine running in the background. Should there be any issue

regarding to the morphometry data files, and error message would be generated and displayed on the screen wall system.

During the multi-level generation of the neuron morphometry data, multiple brains, neurons, annotators are involved. Also, for each neuron, the L1, L2, and L3 data, together with a number of intermediate data levels are produced. The *MorphoHub*-DBMS maintains the correct working state for all the neurons and supports a number of operations, including commit (submitting a reconstruction to a higher level), rollback (sending a reconstruction to a previous level), proofreading (requesting for quality check of the reconstruction), etc. Besides, the *MorphoHub*-DBMS also provides supports for user management and task assignment. All the data are periodically synchronized with private or public cloud-based storages for version control, data sharing and collaboration.

A screen wall display system is connected with *MorphoHub* to present useful information for the ongoing morphometry generation. On the display array, some of the latest reconstructed neurons are displayed; the status of the data repository is tracked; preliminary analysis results are also generated dynamically to provide potential insights of the data.

Generation of Somata and Dendritic Arbors (DendriteGenerator)

The annotators browsed each whole brain image in *D62* using TeraFly and pinpointed soma locations. 2-D maximum intensity projection (MIP) images were generated for each soma-centered region for further validation. Then, centering at each identified soma, a local image volume of $1024 \times 1024 \times 512$ was cropped from the highest resolution of whole brain image. Next, the APP2 (Xiao & Peng, 2013) algorithm was invoked for the tracing of dendritic arbors. In particular, a number of background thresholds (15, 20, ..., 40) were adopted for each tracing routine. The tasks were submitted to the cluster server for parallel computation. The results were retrieved and stored only if the execution time was under 30 s. Then, we leveraged the gold-standard datasets, e.g., a set of manually annotated and validated dendritic arbors, to form rules for further screening the automatic reconstruction results. The [min, max] interval of the following five features of the dendritic arbors were considered, including 'Tips', 'Length', 'Max Path Distance', 'Average Bifurcation Angle Remote', and 'Max Branch Order'. An automatic tracing result qualified if more than four features conformed with the gold-standard. In case more than one tracing qualified for a certain soma location, the result with larger overall tracing length was selected. In visual screening, tracing results were removed when multiple cells were connected.

Generation of L1 and L2 Data

In *MorphoHub*, we use a multi-level approach to reconstruct neuron morphology: the L1 reconstruction contains the dendritic trees and the long axonal projections, while the structures of distal axonal arbors are annotated in the L2 reconstruction. The generation of either L1 or L2 data requires several rounds of checking and correction and is essentially an iterative procedure. In each iteration, there is a generation step (GS) and a validation step (VS) (Supplementary Fig. 7). In the generation step, an annotator tries to reconstruct the neuron's morphology until the person considers that the reconstruction meets the standard of the current level, i.e., L1 or L2. Then, in the validation step, a second annotator examines the reconstruction while labeling the over-traced structures (false positive, FP) and missing structures (false negative, FN) and confirming the correct structures (true positive, TP). After that, the precision rate ($P = TP / (TP + FP)$) and recall rate ($R = TP / (TP + FN)$) can be calculated. If the F1 score = $(2 * RP / (R + P))$ is greater than the preset threshold, the current level is considered finalized. Otherwise, another iteration is needed. Normally, the reconstructions of both L1 and L2 converge in two iterations.

Calculation of the Overall Reconstruction Accuracy in Random-Sampling Simulation

The reconstruction of a neuron's morphology, normally starting from the soma and gradually extending all the way to the remote terminals, is essentially a consecutive decision-making process. Thus, it is likely to suffer from the problem of error propagation, i.e., the reconstruction errors at upstream structures will affect the downstream structures. We assume that for each primitive structure in neuron reconstruction, e.g., a neuronal segment, there is a constant probability p for error occurrence. Besides, if error occurs at some structure, it will also propagate to all its child structures. During the validation step, reconstruction errors could be identified and corrected. However, new errors would still be likely to be introduced at the given chance. Based on such assumptions, we could calculate the overall reconstruction accuracy after n rounds of iterations. With the L1-L2 leveled protocol, we assign $n/2$ iterations for L1 reconstructions, and other $n/2$ iterations for L2. With the one-level strategy, we simply repeat the generation-validation steps n times for the reconstruction of whole neuron. The simulation can be carried out several times to achieve stable results.

L3 (Bouton) Generation (BoutonGenerator)

As a typical L3 data, the bouton distributions are generated based on the guidance of L2 axonal arbors. Boutons are mainly located at distal axonal arbors rather than the long

axonal projections. We used a four-step approach to extract L3 (bouton) data automatically. Firstly, the neuronal tree representation of L2 data is resampled using a fixed-length interval. In this work, the interval is set to be 4 microns. Secondly, for each node in the neuronal tree, the corresponding image intensity value is retrieved from the whole brain datasets. Since the nodes cannot be guaranteed to locate at the centers of the putative boutons, the nodes are allowed to be locally shifted to the maximum intensity position within a small area (e.g., 2 voxels). We assume that the intensity of imaging signal along axons obeys the 1-D Gaussian distribution and a bouton site tends to have intensity jump compared with its neighboring nodes. Thus, the third step is to calculate the intensity jump threshold for each small image block (e.g., $128 \times 128 \times 123$) as the standard deviation of the block, and extract bouton candidate in a block-wise manner. Finally, in the last step, we remove any possibly duplicated bouton site if it is too close to its neighboring site (e.g., within 5 voxels).

Generation of the L0 Imaging Data (L0Generator)

The L0 imaging data is generated based on the corresponding morphometry data, e.g., L1 data, L2 data, or even a dendritic arbor. The L0 data contains the image regions that cover all the anatomical structures of the morphometry and is organized in a TeraFly-compatible hierarchical form, just as the whole-brain imaging data. The approach to generate L0 data is described below. For a given neuron, each fundamental morphological element, i.e., nodes and edges, is examined. The local image block in the whole-brain dataset that contains the element is found and marked as “relevant”. Then, all the “relevant” images are combined into the union set from which the compact L0 data are finally generated by building a multi-resolution image hierarchy.

Quality Control Workflow and Public Release

The reconstructions are concurred to be released after following the single-tree criterion, which includes correct types, no loops, and no trifurcation or multi-furcation. Combination of manual modifications and automatic algorithms were used to control the quality of reconstructions. Automatic routines were invoked to detect any presence of gaps, loop and multi-furcation, followed by manual correction of such issues. Other procedures include examining the reconstruction quality at branch terminals and checking whether all the neurites are centered at image signals. After 2 to 3 rounds of checking, the reconstructions are then ready for ingestion and mirroring in open-source repositories such as NeuroMorpho.Org (<http://neuromorpho.org/>).

Results

We attempted this petabyte (PB)-scale whole-brain computing challenge by introducing a technology that involves a hardware platform that is able to handle petabytes of data storage, sharing, computing and visualization, a software platform called *MorphoHub* that can utilize such hardware platform, and most importantly, a mechanism to scale up the synergized automated computation and multi-user collaboration for effective validation and correction. Our method is centered around reconstructing the multi-level information of single neuron’s morphology (Fig. 1a) at the whole-brain scale for a number of brains. The *MorphoHub* software package is able to streamline the workflow of imaging data management, visualization, reconstruction and collaboration, and data sharing (Fig. 1b, “**Methods**”). Using this approach, we extended several state-of-the-art methods to the PB-scale (Supplementary Table 1) and produced multi-morphometry data from such massive image database (Fig. 1c). Our method allows a smooth transition from manual and interactive morphometry acquisition to increasingly routine work done by automatic algorithms as we show below.

A key component of our method is a three-level (L1, L2, and L3) reconstruction approach (Fig. 1a). We incrementally reconstruct morphological components of neurons, including somata, dendrites, axons, spines and boutons, *only* when such information can be produced faithfully and affordably. Specifically, an L1 reconstruction contains the full dendritic arbor and the skeleton of all axonal neurite tracts, excluding the fine structures of distal axonal arbors (Figs. 1a and 2). An L2 reconstruction contains the complete structures of all neuronal arbors (Figs. 1a and 2). An L3 reconstruction contains the identification of two key elements of synaptic connectivity, dendritic spines and axonal boutons, as well as other structures of potential interest (e.g., specific topology of axonal branching patterns, modeling of specific neuronal compartments’ shape) (Figs. 1a and 3).

The proposed multi-level reconstruction method is generic and scalable to single neuron datasets of arbitrary size if proper data structure and data workflow are in place. To provide such capability for a real PB-scale computing environment, we developed *MorphoHub* to manage all data flow and processing procedures in an integrated way (Fig. 1b and Supplementary Fig. 1, “**Methods**”). *MorphoHub* handles four heterogeneous data types, including image volumes, neuron morphology, meta-data of user interactions, and data management (conversion, storage, transferring/sharing) schemas, for a PB-scale database. We also engineered a universal application interface (Fig. 1b and Supplementary Fig. 2) in *MorphoHub* so that it could invoke additional image analysis and validation tools when needed.

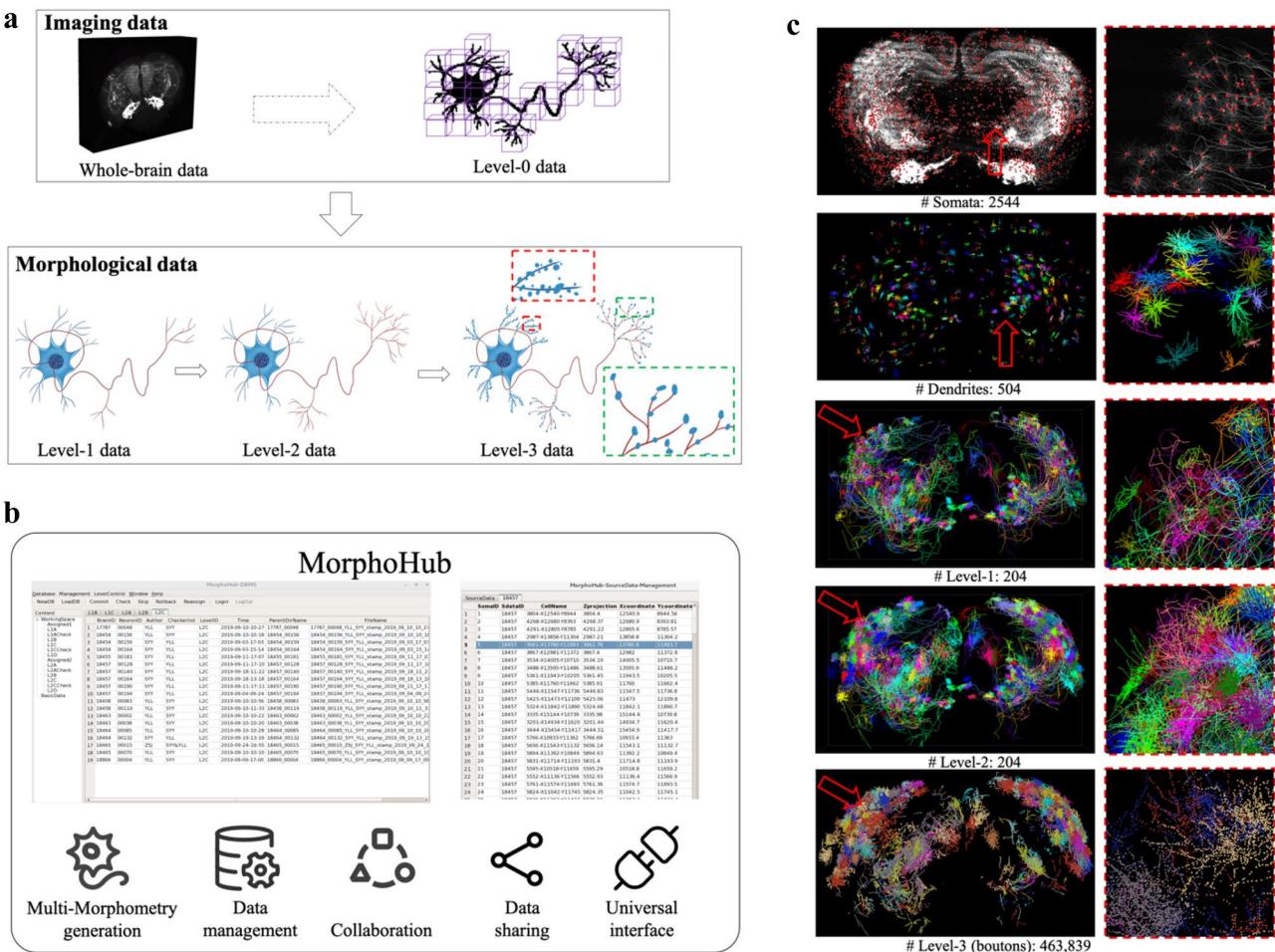


Fig. 1 Multi-morphometry data generation from whole-brain imaging datasets **a** An illustration of the multi-level reconstruction approach. From a whole-brain image containing trillions of voxels (top left), the Level-1 (L1), Level-2 (L2), and Level-3 (L3) data are reconstructed in sequence (bottom). Moreover, a concise Level-0 (L0) imaging data is also generated based on the reconstructed morphometry (top right) **b**

The MorphoHub system for the generation of multi-morphometry data, management and visualization of all related data and workflow, data sharing and extended functions **c** Examples of the multi-morphometry data reconstructed from one Brain (Brain id: 18,454). From top to bottom are the somata, dendrites, L1, L2, and L3 data, respectively, with zoom-in panels for red arrows shown on the right

To demonstrate the capability of this approach, we built an image database called *D62* consisting of 62 whole mouse brain images. *D62* has in total 713.35 teravoxels, 1.43 petabytes in native image space, and 973 terabytes in compressed file space (Supplementary Data 1). *MorphoHub* ran smoothly on *D62* and allowed us to precisely pinpoint somata of 50,233 neurons using TeraFly (100% accuracy validated by independent annotators). For each neuron, we then reconstructed the dendrites automatically (Fig. 3a, “Methods”), followed by feature-based screening and visual validation (Fig. 1c, “Methods”). Using this workflow, we produced traceable dendritic results of 11,322 neurons. Due to the scale of the problem, similar results were hard to obtain using other software.

For each of the sparsely labeled neurons whose long axon projection could be separated, we first produced an

L1-reconstruction corresponding to the key skeleton of a neuron along with its dendrites and axonal projection targets. We then requested human annotators to validate each L1-reconstruction. The resulting L1-reconstruction was then further refined to complete the L2-reconstruction that also added the distal axonal arbors projecting far away across various brain regions. In this study, we focused on reconstructing the full morphology of 1,050 neurons whose somata situate in the thalamus, striatum, and cortical regions of mouse brains (“Methods”). Each L1-L2 pair of the completed neuron reconstructions were validated by at least two annotators. This dataset, called *R1050*, was used to further examine whether or not the core multi-level reconstruction method would make sense.

We compared the L1 and L2 morphology in *R1050*. On average a pair of L1-L2 reconstructions have five to six fold

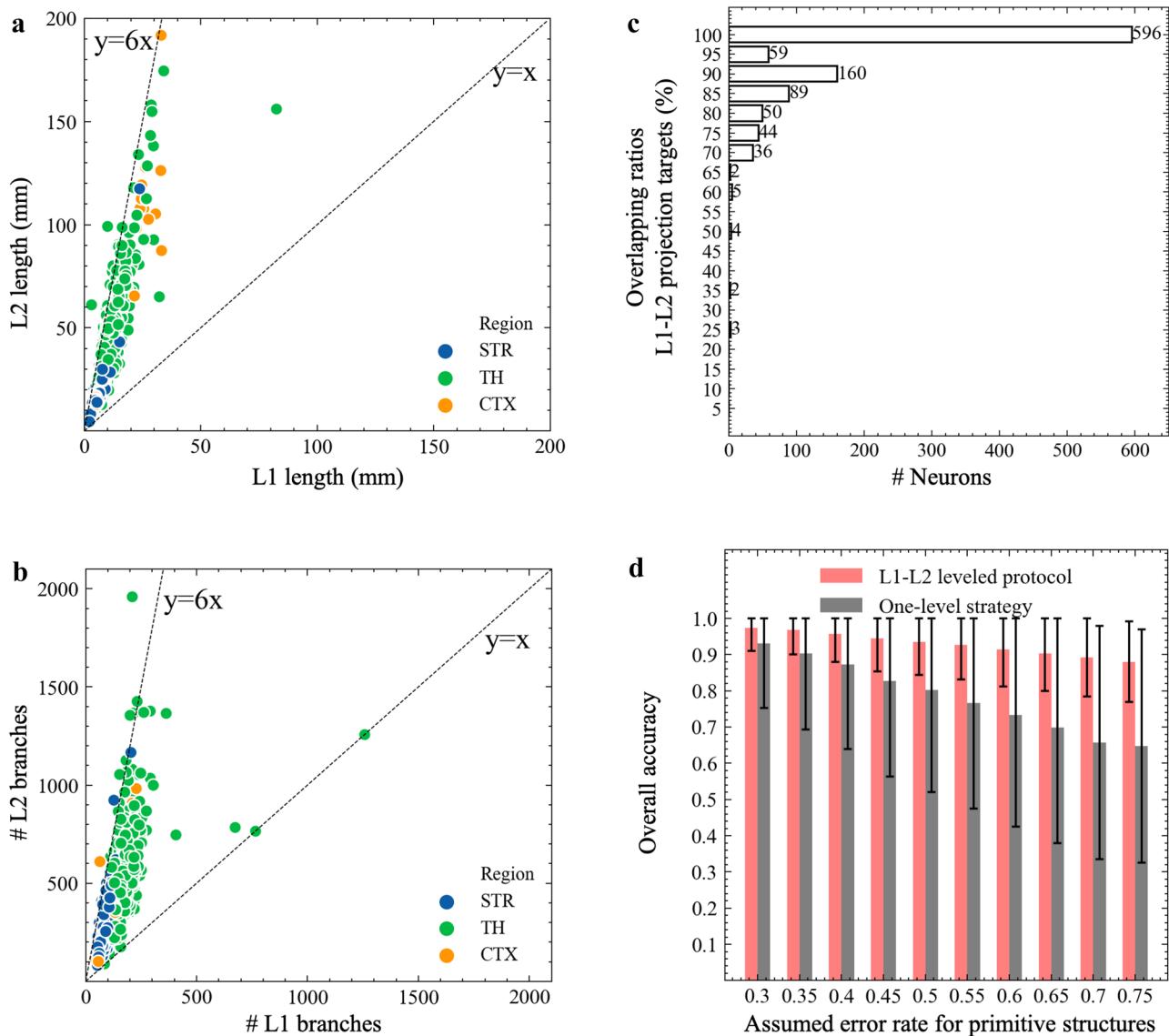


Fig. 2 Comparisons of L1 and L2 data. **a** and **b** Two features (**a**: total length; **b**: branch number) of L1 and L2 data. Neurons are from three brain regions (TH (Thalamus): 705 neurons; CTX (Cortex): 23 neurons; STR (Striatum): 322 neurons). **c** Comparison of the projection targets of the L1-L2 pairs. Horizontal axis: number of neurons. Vertical axis: overlapping ratios of projection targets of L1-L2 pairs. **d**

Random-sampling simulation of potential reconstruction errors of the L1-L2 leveled protocol and the brute-force one-level strategy. Horizontal axis: the assumed error rate for each primitive structure (e.g., a tract). Vertical axis: overall accuracy of the reconstructions after validation. Error bar: SD

difference in terms of their length and number of branches (Fig. 2a and b). However, a Sholl analysis (Langhammer et al., 2010) indicates that the L1-L2 pairs share branching patterns in dendrites and differ only in the additional axonal regions of L2 (Supplementary Fig. 3). Additionally, 95% of the corresponding L1 and L2 reconstruction-pairs have at least 75% overlap in their projecting target brain-regions (Fig. 2c). Finally, we used random-sampling to simulate the potential reconstruction error that would be seen if the L1-L2 leveled protocol were not used (“Methods”). The L1-L2 reconstruction protocol can avoid reconstruction

error of a brute-force approach that would reconstruct distal arbors directly without first validating the L1-reconstruction (Fig. 2d). In summary, Fig. 2 shows that reconstructing L1 first without incurring the great complexity in producing the L2 data already offers an efficient way to analyze the key branching structures and the ballpark projection pattern of single neurons. The reconstructions in each level state were also transmitted and stored efficiently and safely.

On the other hand, the rich detail in the axonal arbors of the L2 reconstruction of a neuron as demonstrated in Fig. 2 also serves as the basis of finer-resolution morphometry.

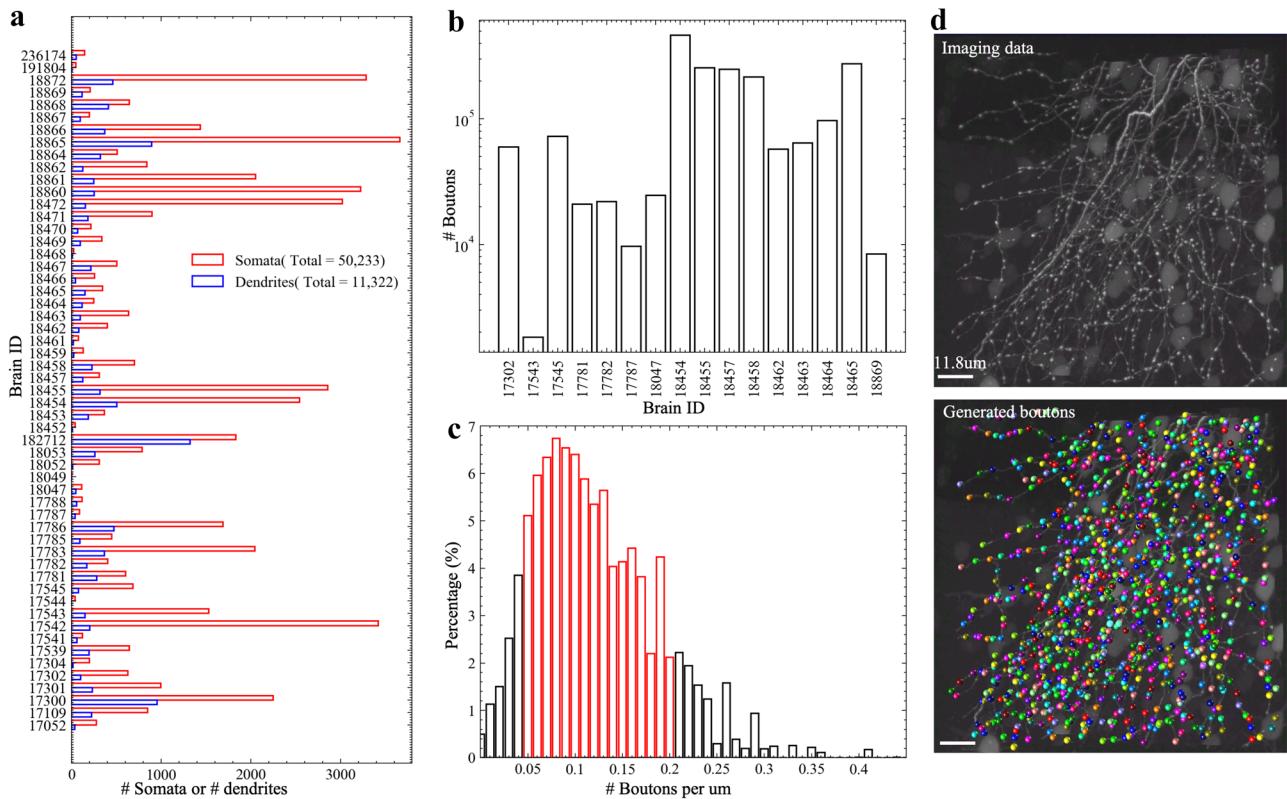


Fig. 3 Mass data production of somata, dendrites, and boutons. **a** Number of identified somata and automatically reconstructed dendritic arbors from D62. **b** Number of detected boutons based on the L2 morphometry in R1050. **c** Distribution of the linear density of boutons along the axonal skeletons. Red bar highlights the range [0.05, 0.2]. **d** Examples of detected boutons. Upper: a local image region containing an axonal cluster. Bottom: putative boutons shown in distinct colors.

tons along the axonal skeletons. Red bar highlights the range [0.05, 0.2]. **d** Examples of detected boutons. Upper: a local image region containing an axonal cluster. Bottom: putative boutons shown in distinct colors.

For axonal arbors in *R1050*, we further generated the L3 morphometry by detecting putative synaptic sites (boutons) of neurons at very large scale (“Methods”). Due to the presence of axonal fibers, we are able to reduce the conventionally required 3-D blob-segmentation problem for bouton detection to a much simpler 1-D Gaussian model fitting problem, that can be solved in both high accuracy and high speed. We fit a Gaussian kernel to the axonal varicosities and detected 1.9 million boutons (Fig. 3b-d). Random inspection of the results indicated that the detection precision was above 95% (Supplementary Fig. 4). We also found that nearly 80% of the average spacing of adjacent boutons along an axon ranged from 5 to 20 μm (Fig. 3c). The L1-L2-L3 trio morphometry produced using this method allows studying the complete distribution of single neurons, their projection, and potential connectivity patterns at whole brain scale. A detailed analysis of the patterns in and statistics of these datasets will be reported elsewhere.

All multi-morphometry data produced were also registered to the Allen Common Coordinate Framework (CCF) (Wang et al., 2020) to see how the distributions of each data level correlate with others (Fig. 4). In this way, various brain regions (the white colored regions), dendrites, axons, and

boutons can all be identified (Fig. 4a-e). As a result, we output a summary matrix with rows representing neurons, columns representing a unique CCF parcel, and numerical entries expressing, for each neuron and corresponding parcel, the axonal and dendritic length, the number of boutons, and the (binary) presence of soma. Such representation lends itself to highly informative quantitative analyses, such as pairwise probability of directional connection between neurons (dot product of presynaptic axonal values and postsynaptic dendritic values) and projection similarity (arccosine distance between axonal values of two neurons). Figure 4 shows that the distributions of these neuronal entities do not correlate globally. Instead, they exhibit regional enrichment of which the pattern is hard to observe when only local brain areas are analyzed.

For PB-scale computing, the speed of data I/O for storage and data sharing across networks (Internet or intranets) becomes more critical than in applications at small scale. It is essential to reduce data volume without compromising visualization and analysis of such large data. We observed that an L1-L2-L3 morphology trio of a neuron will always be sparse and that the spine and bouton information in the L3 data could be described using a neighborhood around the

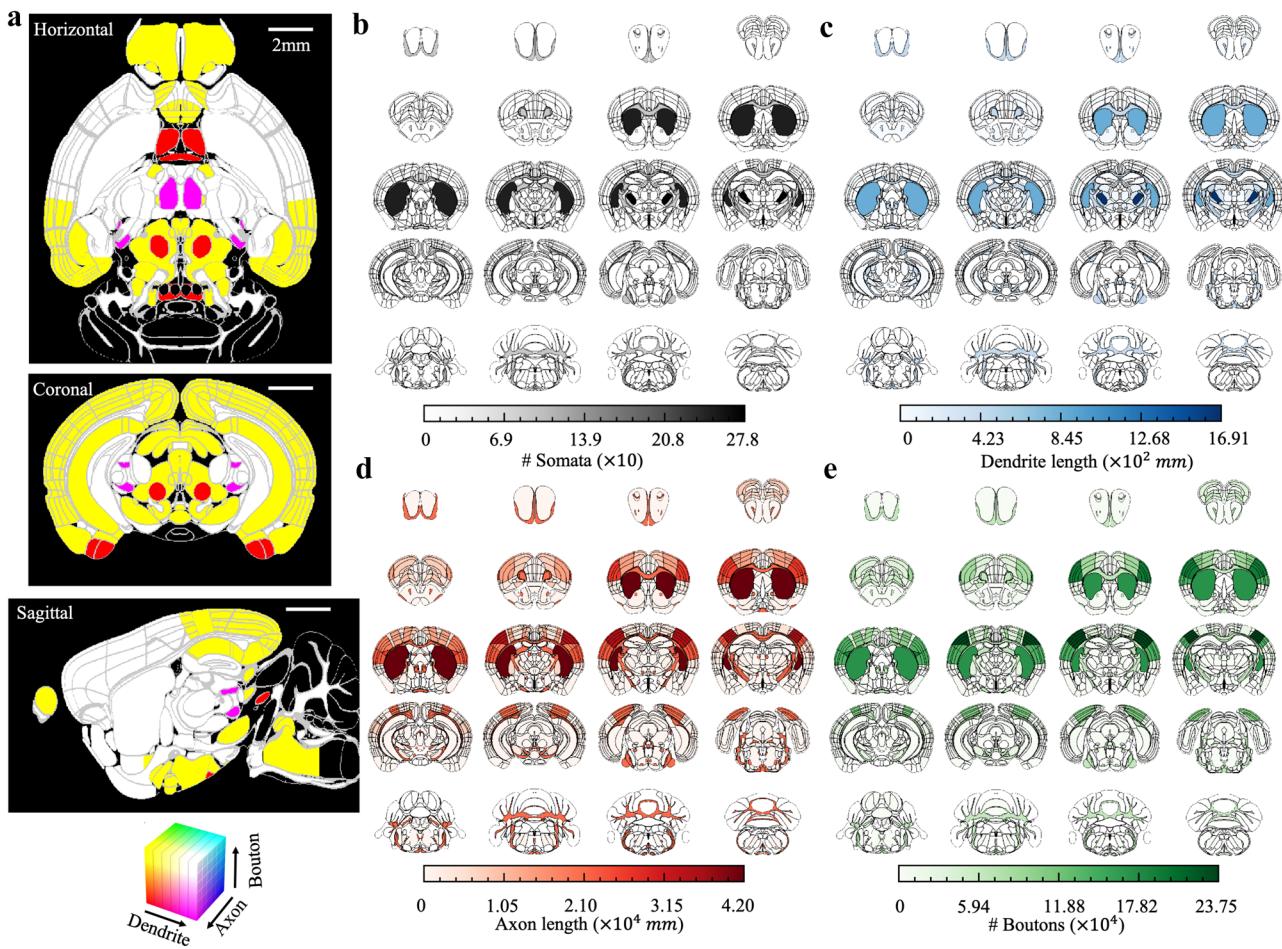


Fig. 4 Visualization of the multi-morphometry data in R1050. **a** The color-coded joint distribution of dendrites, axons, and boutons. From top to bottom: horizontal view (slice no. 165), coronal view (slice no. 335), and sagittal view (slice no. 148). Colors indicate the densities of dendrites, axons, and boutons normalized to the standard RGB color space. Scale bar: 2 mm. **b** Individual distributions of somata in

R1050. Each inset corresponds to the combination of 25 consecutive coronal slices, in which the brain regions were colored according to the densities of somata. The darker the color, the higher the density. **c–e** Similar visualizations for dendrites, axons, and boutons in R1050, respectively

neuron skeleton in preceding levels. To utilize this observation, we developed a compact L0 representation of a neuron for effective imaging data management, sharing, and computation (Fig. 5a, “Methods”). The key idea is that the L0 data of a neuron represents a tightly bounded image region that covers the L1-L2-L3 trio area. Because dendritic spines typically attach dendritic fiber orthogonally, and axonal boutons scatter along axons, for any given L2 (or L1) skeleton we conveniently extracted a series of piecewise $512 \times 512 \times 512$ -voxel 3-D image-tiles around the skeleton to cover all spines and boutons and at the same time to allow fast image file I/O. In this way, an L0 representation identifies an image sub-region that contains all parts of a specific neuron.

For R1050 reconstructed from D62, on average the L0 data generated based on an L1 skeleton contains more than 77% of that generated based on the corresponding L2 reconstruction (Fig. 5b). In addition, the L0 data of a neuron

typically occupies three orders of magnitude less image volume compared to the whole brain image (Fig. 5c). The L0 data of the largest neuron in this work has ~ 80 gigavoxels, while the mean value and standard deviation of the volume of L0 data of 1,050 fully reconstructed neurons are 6.75 and 5.94 gigavoxels, respectively (Fig. 5d, Supplementary Table 2). Practically, even the union of all L0 data of neurons, denoted as Super L0, in a sparsely labeled brain still has 1 to 3 orders of magnitude fewer voxels compared to the total volume of the brain (Fig. 5c). In this way, the multi-morphometry framework allows thousands of fold better efficiency in both storing and transferring the essential image data and quantitative shape information of neurons. This utility greatly simplifies the previously challenging data sharing task. Indeed, without accelerated content delivery, currently it is possible to transfer the L0 data in R1050 between the data production center (SEU-ALLEN) in Asia

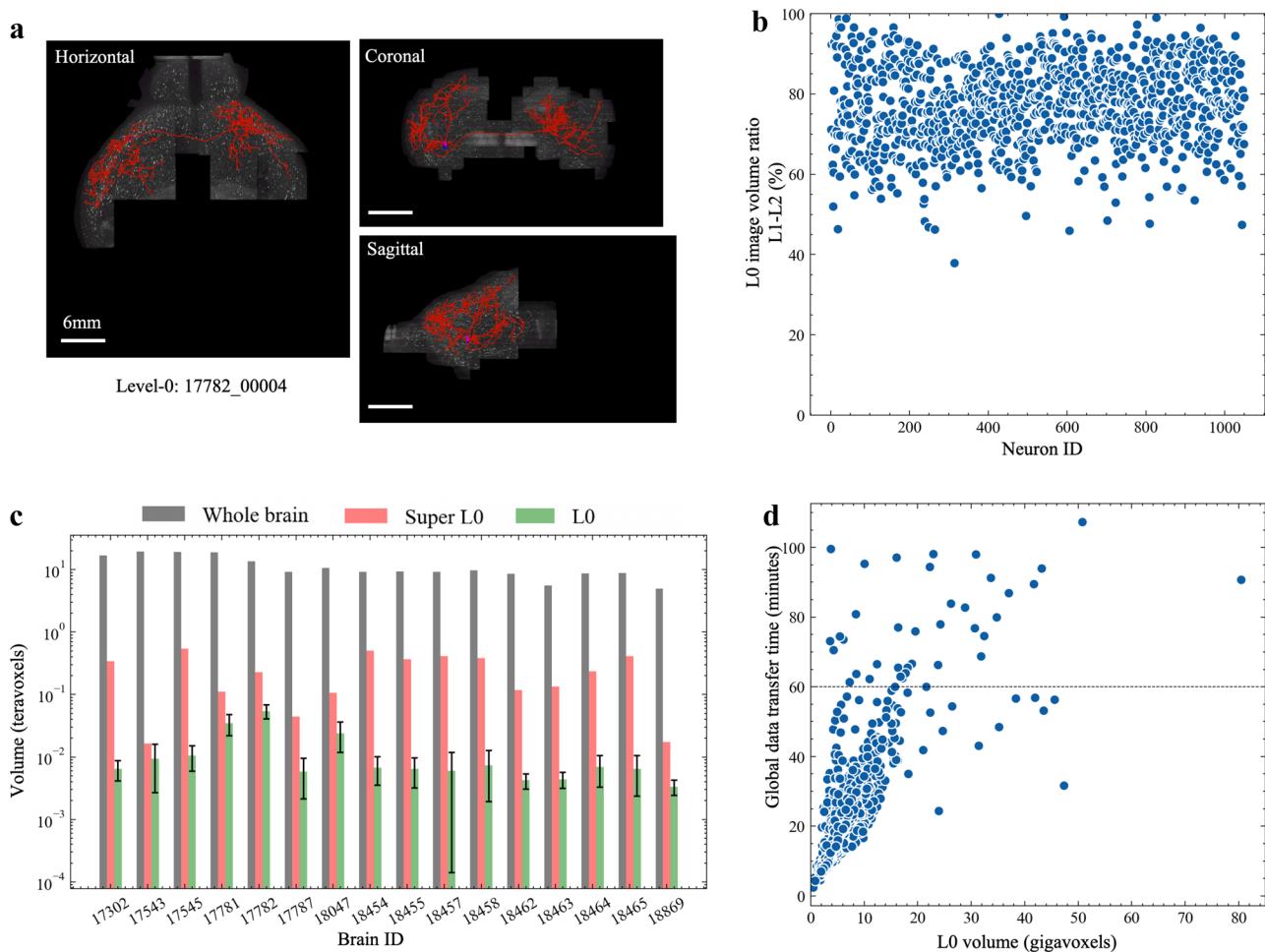


Fig. 5 The L0 representation of imaging data **a** The L0 image (shown in horizontal, coronal, and sagittal views) for neuron 17782_00004, overlaid with its L2 reconstruction **b** For neurons in R1050, the ratio of L0 image volume generated from L1 data over that generated from L2 data **c** Comparisons of the size of whole-brain images, the aver-

age size of L0 data, and the size of the “super L0-data” (union of all L0 data of neurons). Error bar: SD **d** Time for transferring 1050 L0 images between two research centers in Asia (SEU-ALLEN) and America (BIL)

and one data releasing facility (BICCN Image Library, Pittsburgh supercomputing center) in North America (Fig. 5d). This direct data sharing replaced a previous way to bulk ship hard drives containing the massive amount data back-and-forth across continents.

With the L0-L1-L2-L3 quadruple data, we further enhanced the scale and faithfulness of our multi-morphometry produced in two ways. First, since the L0 data has a much smaller volume and thus is much easier to share across network, we developed a collaborative mode that interconnects a number of formerly autonomous TeraFly/TeraVR users to synergistically work on the L0 data directly (Supplementary Fig. 5). This method parallelizes the workflow and thus improves the data production rate. The cooperative work of multiple annotators also elevates the faithfulness of the resultant morphometry. Second, we used a deep learning network to learn from the L0-Lx ($x=1,2,3$) pair. The trained

model was used to detect specific neuronal features, such as neurite skeleton or axonal terminals (Supplementary Fig. 6). This process can be repeatedly optimized based on progressively more and more accurate L0-L1-L2-L3 quadruple data. Such automation also increases the data production rate for PB-scale computing.

Discussion

In this study, we demonstrated a robust PB-scale informatics platform to generate large-scale single neuron reconstructions suitable for multi-scale biological analysis. Our approach has several advantages: (1) efficient multi-level production and management of whole-brain neuron reconstructions; (2) conducting morphological analysis and cell typing globally and at multi-resolution; (3) enabling the investigation of the

convergence or divergence of neuronal projections by analyzing distribution of neuronal arbors across brain regions; (4) comparison of various neuronal elements and substructures with respect to the types of cells. Taken altogether, our whole-brain multi-morphometry approach provides a framework to produce hierarchical datasets that synchronize brain anatomy, single neuron morphology, sub-neuronal structures, and potential pre-synaptic sites, all mapped onto a standardized atlas. Our method will be useful for further studies of neuronal circuits based on whole-brain imaging, not only for mouse brains but also for other model systems such as monkey brains.

Our work furthers previous effort to use light microscopy (LM) to visualize and detect synapses around neurite tracts labeled by genetic markers (Iascone et al., 2020; Kim et al., 2012) or antibodies (Micheva & Smith, 2007), which were limited to partial neuronal structures in local brain regions. In this study we used fMOST data as a showcase and for putative synaptic sites we have focused on axonal boutons. As a generic computational framework, our approach is applicable to various datasets produced with different methods and collected with different imaging modalities.

Of note, our method stores both the original neuronal tracing in the native coordinates of the individual brain specimen, allowing efficient extraction of precise geometric measurements (e.g., synaptic distance from the soma along the axonal path) as well as a registered version of the same morphology mapped to CCF. The registration of the axonal reconstructions is efficiently expressed by augmenting each tracing coordinate, from the center of the soma all the way to the boutons, with the unique identifier of the anatomical parcel in which it is embedded (Nanda et al., 2018). On the one hand, such compact representation immediately enables real-time computation of potential circuit connectivity (Rees et al., 2017). On the other, it provides critical information regarding neuronal identity by encoding its somatic region and quantitative projection targets. This information, together with the specification of essential details regarding the brain specimen and imaging modality, fulfills the recommended requirements for standard metadata description of neuromorphological reconstructions (Bijari et al., 2020). Thus, our IT infrastructure is seamlessly compliant for effective pipelining with the NeuroMorpho.Org public data sharing platform, ensuring maximum impact through expanded community outreach (Ascoli et al., 2017).

Another line of major current effort is to reconstruct dense neurite tracts and synapses, or “connectomes”, using electron microscopy (EM) followed by computationally intensive 3D image segmentation and modeling, such as the MICRON project (Yin et al., 2020) and the Janelia FlyEM project (Xu et al., 2020). For mammalian brains, EM-based approaches are still restricted to local brain regions due to their very high resolution and various challenges in sample

preparation and imaging. Our method complements the EM-based approach in whole brain scale profiling. Importantly, the lower cost of LM enables one to integrate morphometry information from many brains that is crucial to understand the variability of neurons and their circuits across brains and conditions. In comparison with EM-based reconstruction of individual neurons, our work reconstructs the whole morphology, including (1) soma locations; (2) dendritic arbors, which entails the capacity for and locations of synaptic inputs; (3) axonal trajectories with collateral projections and terminal boutons, which indicate innervation of projection targets. Together, such information is directly relevant to neuronal function.

The multi-level reconstruction approach is being enhanced in various ways. In addition to various workstation/PC clients, virtual reality consoles, super-computing, and big-display walls that are already integrated in our software *MorphoHub*, mobile applications (APPs) for more intelligent and automated neuron tracing are being developed and added onto our software. We are also deploying *MorphoHub* for data servers in the cloud and scaling up the capability for concurrent data serving of distributed users. We hope these engineering efforts would lead to a new globally accessible platform that has potential to bring the current productivity to the next level, especially addressing challenges in completing neuron morphometry more efficiently, producing more fine-scale morphometry such as synapses with their shapes and statistics, integrating more automation through the use of AI, sharing of imaging data remotely at affordable cost, and international collaboration of neuroanatomists and other interested users.

Conclusion

Neuronal morphology is an essential component of cell type identity in the brain and an essential determinant of connectivity and circuit function. Large scale accurate neuronal profiling necessitates advanced methods in computational processing to effectively manage storage and bandwidth for collaborative segmentation and annotation. The data in this study shows our petabyte-scale computing framework is able to provide a solution to modern anatomic workflow requirements that are now demanded for very large-scale morphometry.

Information Sharing Statement

The whole brain image datasets are released under BICCN’s Brain Image Library (BIL) at Pittsburgh Supercomputing Center. The multi-morphometry datasets can be downloaded at <https://github.com/SD-Jiang/MorphoHub/releases/tag/>

v1.0. The software is available at <https://github.com/SD-Jiang/MorphoHub>.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12021-022-09569-4>.

Acknowledgements This work is funded by Southeast University (SEU). SEU supported the development of the methods and informatics data management and analysis pipeline at the SEU-Allen Joint Center. We are grateful to the Allen Institute for imaging datasets collected through multiple grant awards from institutes under the National Institutes of Health (NIH), including award number R01EY023173 from The National Eye Institute to H.Z., U01MH105982 from the National Institute of Mental Health and Eunice Kennedy Shriver National Institute of Child Health & Human Development to H.Z., and U19MH114830 from the National Institute Of Mental Health to H.Z. Y.W. acknowledges National Science Foundation of China Grant number 32071367 and Natural Science Foundation Shanghai Grant 20ZR1420100. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH and its subsidiary institutes. The Allen Institute affiliated authors wish to thank the Allen Institute founder, Paul G. Allen, for his vision, encouragement, and support. G.A.A. acknowledges NIH grants R01NS39600, U01MH14829, and R01NS86082. We thank Zhi Zhou, Yuanyuan Song, Lulu Yin, Shichen Zhang, Jintao Pan, Yanting Liu, Guodong Hong, Jia Yuan, Yanjun Duan, Yaping Wang, Qiang Ouyang, Zijun Zhao, Wan Wan, Peng Wang, Ping He, Lingsheng Kong, Feng Xiong, and other team members for the support of data production.

Authors' Contributions H.P. conceptualized the study, envisioned and led the development of this platform and various analyses. S.J., Y.W., and L.L. co-developed the *MorphoHub* system and the multi-level neuron reconstruction protocols. L.L. led the reconstruction of *R1050*. Y.W. provided the support of *TeraVR*. M.C. led the identification of somata in D62. S.Z., X.Z. assisted in *MorphoHub* development. P.X. and L.D. assisted in data analysis. Z.R. and H.P. constructed the hardware platform for this study. H.Z. provided the raw imaging dataset for producing D62 and advised on neurobiology. M.H. advised on morphometry and data management. H.D. and G.A. advised on whole brain neuronal structure analysis. H.P. led the writing of the manuscript in consultation with all authors.

Declarations

Conflicts of Interest The authors declare no competing interests.

References

- Acciai, L., Soda, P., & Iannello, G. (2016). Automated Neuron Tracing Methods: An Updated Account. *Neuroinformatics*, 14(4), 353–367. <https://doi.org/10.1007/s12021-016-9310-0>
- Aransay, A., Rodríguez-López, C., García-Amado, M., Clascá, F., & Prensa, L. (2015). Long-range projection neurons of the mouse ventral tegmental area: A single-cell axon tracing analysis. *Frontiers in Neuroanatomy*, 9, 59. <https://doi.org/10.3389/fnana.2015.00059>
- Ascoli, G. A., Maraver, P., Nanda, S., Polavaram, S., & Armañanzas, R. (2017). Win-win data sharing in neuroscience. *Nature Methods*, 14(2), 112–116. <https://doi.org/10.1038/nmeth.4152>
- Bass, C., Helkkula, P., Paola, V. D., Clopath, C., & Bharath, A. A. (2017). Detection of axonal synapses in 3D two-photon images. *PLoS One*, 12(9), e0183309. <https://doi.org/10.1371/journal.pone.0183309>
- Bijari, K., Akram, M. A., & Ascoli, G. A. (2020). An open-source framework for neuroscience metadata management applied to digital reconstructions of neuronal morphology. *Brain Informatics*, 7(1), 2. <https://doi.org/10.1186/s40708-020-00103-3>
- BRAIN Initiative Cell Census Network (BICCN). (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature*, 598, 86–102. <https://doi.org/10.1038/s41586-021-03950-0>
- Bria, A., Iannello, G., Onofri, L., & Peng, H. (2016). TeraFly: Real-time three-dimensional visualization and annotation of terabytes of multidimensional volumetric images. *Nature Methods*, 13(3), 192–194. <https://doi.org/10.1038/nmeth.3767>
- Cheng, S., Wang, X., Liu, Y., Su, L., Quan, T., Li, N., Yin, F., Xiong, F., Liu, X., Luo, Q., Gong, H., & Zeng, S. (2019). DeepBouton: Automated Identification of Single-Neuron Axonal Boutons at the Brain-Wide Scale. *Frontiers in Neuroinformatics*, 13, 25. <https://doi.org/10.3389/fninf.2019.00025>
- Cohen, A. R., Roysam, B., & Turner, J. N. (1994). Automated tracing and volume measurements of neurons from 3-D confocal fluorescence microscopy data. *Journal of Microscopy*, 173(2), 103–114. <https://doi.org/10.1111/j.1365-2818.1994.tb03433.x>
- Economou, M. N., Clack, N. G., Lavis, L. D., Gerfen, C. R., Svoboda, K., Myers, E. W., & Chandrashekhar, J. (2016). A platform for brain-wide imaging and reconstruction of individual neurons. *eLife*, 5, e10566. <https://doi.org/10.7554/eLife.10566>
- Eliceiri, K. W., Berthold, M. R., Goldberg, I. G., Ibáñez, L., Manjunath, B. S., Martone, M. E., Murphy, R. F., Peng, H., Plant, A. L., Roysam, B., Stuurman, N., Swedlow, J. R., Tomancak, P., & Carpenter, A. E. (2012). Biological imaging software tools. *Nature Methods*, 9(7), 697–710. <https://doi.org/10.1038/nmeth.2084>
- Gala, R., Lebrecht, D., Sahlender, D. A., Jorstad, A., Knott, G., Holtmaat, A., & Stepanys, A. (2017). Computer assisted detection of axonal bouton structural plasticity in vivo time-lapse images. *eLife*, 6, e29315. <https://doi.org/10.7554/eLife.29315>
- Gillette, T., Brown, K., Svoboda, K., Liu, Y., & Ascoli, G. (2011). DIADEMchallenge.Org: A Compendium of Resources Fostering the Continuous Development of Automated Neuronal Reconstruction. *Neuroinformatics*, 9, 303–304. <https://doi.org/10.1007/s12021-011-9104-3>
- Gong, H., Xu, D., Yuan, J., Li, X., Guo, C., Peng, J., Li, Y., Schwarz, L. A., Li, A., Hu, B., Xiong, B., Sun, Q., Zhang, Y., Liu, J., Zhong, Q., Xu, T., Zeng, S., & Luo, Q. (2016). High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. *Nature Communications*, 7(1), 12142. <https://doi.org/10.1038/ncomms12142>
- Gong, H., Zeng, S., Yan, C., Lv, X., Yang, Z., Xu, T., Feng, Z., Ding, W., Qi, X., Li, A., Wu, J., & Luo, Q. (2013). Continuously tracing brain-wide long-distance axonal projections in mice at a one-micron voxel resolution. *NeuroImage*, 74, 87–98. <https://doi.org/10.1016/j.neuroimage.2013.02.005>
- Iascone, D. M., Li, Y., Sümbül, U., Doron, M., Chen, H., Andreu, V., Goudy, F., Blockus, H., Abbott, L. F., Segev, I., Peng, H., & Polleux, F. (2020). Whole-Neuron Synaptic Mapping Reveals Spatially Precise Excitatory/Inhibitory Balance Limiting Dendritic and Somatic Spiking. *Neuron*, 106(4), 566–578. <https://doi.org/10.1016/j.neuron.2020.02.015>
- Karube, F., Kubota, Y., & Kawaguchi, Y. (2004). Axon branching and synaptic bouton phenotypes in GABAergic nonpyramidal cell subtypes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 24(12), 2853–2865. <https://doi.org/10.1523/JNEUROSCI.4814-03.2004>
- Kim, J., Zhao, T., Petralia, R. S., Yu, Y., Peng, H., Myers, E., & Magee, J. C. (2012). MGRASP enables mapping mammalian synaptic connectivity with light microscopy. *Nature Methods*, 9(1), 96–102. <https://doi.org/10.1038/nmeth.1784>
- Langhammer, C. G., Previtera, M. L., Sweet, E. S., Sran, S. S., Chen, M., & Firestein, B. L. (2010). Automated Sholl analysis of digitized

- neuronal morphology at multiple scales: Whole cell Sholl analysis versus Sholl analysis of arbor subregions. *Cytometry Part a: The Journal of the International Society for Analytical Cytology*, 77(12), 1160–1168. <https://doi.org/10.1002/cyto.a.20954>
- Liu, M., Chen, W., Wang, C., & Peng, H. (2019). A Multiscale Ray-Shooting Model for Termination Detection of Tree-Like Structures in Biomedical Images. *IEEE Transactions on Medical Imaging*, 38(8), 1923–1934. <https://doi.org/10.1109/TMI.2019.2893117>
- Markram, H. (2006). The Blue Brain Project. *Nature Reviews Neuroscience*, 7(2), 153–160. <https://doi.org/10.1038/nrn1848>
- Meijering, E., Carpenter, A. E., Peng, H., Hamprecht, F. A., & Olivo-Marin, J.-C. (2016). Imagining the future of bioimage analysis. *Nature Biotechnology*, 34(12), 1250–1255. <https://doi.org/10.1038/nbt.3722>
- Micheva, K. D., & Smith, S. J. (2007). Array tomography: A new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron*, 55(1), 25–36. <https://doi.org/10.1016/j.neuron.2007.06.014>
- Myers, G. (2012). Why bioimage informatics matters. *Nature Methods*, 9(7), 659–660. <https://doi.org/10.1038/nmeth.2024>
- Nanda, S., Chen, H., Das, R., Bhattacharjee, S., Cuntz, H., Torben-Nielsen, B., Peng, H., Cox, D. N., De Schutter, E., & Ascoli, G. A. (2018). Design and implementation of multi-signal and time-varying neural reconstructions. *Scientific Data*, 5(1), 170207. <https://doi.org/10.1038/sdata.2017.207>
- Parekh, R., & Ascoli, G. A. (2013). Neuronal Morphology goes Digital: A Research Hub for Cellular and System Neuroscience. *Neuron*, 77(6), 1017–1038. <https://doi.org/10.1016/j.neuron.2013.03.008>
- Peng, H., Bria, A., Zhou, Z., Iannello, G., & Long, F. (2014). Extensible visualization and analysis for multidimensional images using Vaa3D. *Nature Protocols*, 9(1), 193–208. <https://doi.org/10.1038/nprot.2014.011>
- Peng, H., Hawrylycz, M., Roskams, J., Hill, S., Spruston, N., Meijering, E., & Ascoli, G. A. (2015). BigNeuron: Large-scale 3D Neuron Reconstruction from Optical Microscopy Images. *Neuron*, 87(2), 252–256. <https://doi.org/10.1016/j.neuron.2015.06.036>
- Peng, H., Ruan, Z., Atasoy, D., & Sternson, S. (2010). Automatic reconstruction of 3D neuron structures using a graph-augmented deformable model. *Bioinformatics*, 26(12), i38–i46. <https://doi.org/10.1093/bioinformatics/btq212>
- Peng, H., Xie, P., Liu, L., Kuang, X., Wang, Y., Qu, L., Gong, H., Jiang, S., Li, A., Ruan, Z., Ding, L., Chen, C., Chen, M., Daigle, T. L., Ding, Z., Duan, Y., Feiner, A., He, P., Hill, C., & Zeng, H. (2021). Morphological diversity of single neurons in molecularly defined cell types. *Nature*, 598, 174–181. <https://doi.org/10.1038/s41586-021-03941-1>
- Peng, H., Zhou, Z., Meijering, E., Zhao, T., Ascoli, G. A., & Hawrylycz, M. (2017). Automatic tracing of ultra-volumes of neuronal images. *Nature Methods*, 14(4), 332–333. <https://doi.org/10.1038/nmeth.4233>
- Pietzsch, T., Saalfeld, S., Preibisch, S., & Tomancak, P. (2015). BigDataViewer: Visualization and processing for large image data sets. *Nature Methods*, 12(6), 481–483. <https://doi.org/10.1038/nmeth.3392>
- Rees, C. L., Moradi, K., & Ascoli, G. A. (2017). Weighing the Evidence in Peters' Rule: Does Neuronal Morphology Predict Connectivity? *Trends in Neurosciences*, 40(2), 63–71. <https://doi.org/10.1016/j.tins.2016.11.007>
- Ropireddy, D., Scorcioni, R., Lasher, B., Buzsáki, G., & Ascoli, G. A. (2011). Axonal morphometry of hippocampal pyramidal neurons semi-automatically reconstructed after *in vivo* labeling in different CA3 locations. *Brain Structure & Function*, 216(1), 1–15. <https://doi.org/10.1007/s00429-010-0291-8>
- Rotolo, T., Smallwood, P. M., Williams, J., & Nathans, J. (2008). Genetically-Directed, Cell Type-Specific Sparse Labeling for the Analysis of Neuronal Morphology. *PLoS One*, 3(12), e4099. <https://doi.org/10.1371/journal.pone.0004099>
- Sivagnanam, S., Majumdar, A., Yoshimoto, K., Astakhov, V., Bandrowski, A., Martone, M., & Carnevale, N. T. (2015). Early experiences in developing and managing the neuroscience gateway. *Concurrency and Computation: Practice and Experience*, 27(2), 473–488. <https://doi.org/10.1002/cpe.3283>
- Swedlow, J. R., Goldberg, I., Brauner, E., & Sorger, P. K. (2003). Informatics and quantitative analysis in biological imaging. *Science (New York, N.Y.)*, 300(5616), 100–102. <https://doi.org/10.1126/science.1082602>
- Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., Dolbeare, T., Blanchard, B., Dee, N., Wakeman, W., Hirokawa, K. E., Szafer, A., Sunkin, S. M., Oh, S. W., Bernard, A., & Ng, L. (2020). The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell*, 181(4), 936–953. <https://doi.org/10.1016/j.cell.2020.04.007>
- Wang, Y., Li, Q., Liu, L., Zhou, Z., Ruan, Z., Kong, L., Li, Y., Wang, Y., Zhong, N., Chai, R., Luo, X., Guo, Y., Hawrylycz, M., Luo, Q., Gu, Z., Xie, W., Zeng, H., & Peng, H. (2019). TeraVR empowers precise reconstruction of complete 3-D neuronal morphology in the whole brain. *Nature Communications*, 10(1), 3474. <https://doi.org/10.1038/s41467-019-11443-y>
- Winnubst, J., Bas, E., Ferreira, T. A., Wu, Z., Economo, M. N., Edson, P., Arthur, B. J., Bruns, C., Rokicki, K., Schauder, D., Olbris, D. J., Murphy, S. D., Ackerman, D. G., Arshadi, C., Baldwin, P., Blake, R., Elsayed, A., Hasan, M., Ramirez, D., & Chandrashekhar, J. (2019). Reconstruction of 1,000 Projection Neurons Reveals New Cell Types and Organization of Long-Range Connectivity in the Mouse Brain. *Cell*, 179(1), 268–281. <https://doi.org/10.1016/j.cell.2019.07.042>
- Xiao, H., & Peng, H. (2013). APP2: Automatic tracing of 3D neuron morphology based on hierarchical pruning of a gray-weighted image distance-tree. *Bioinformatics (oxford, England)*, 29(11), 1448–1454. <https://doi.org/10.1093/bioinformatics/btt170>
- Xu, C. S., Januszewski, M., Lu, Z., Takemura, S., Hayworth, K. J., Huang, G., Shinomiya, K., Maitin-Shepard, J., Ackerman, D., Berg, S., Blakely, T., Bogovic, J., Clements, J., Dolafit, T., Hubbard, P., Kainmueller, D., Katz, W., Kawase, T., Khairy, K. A., & Plaza, S. M. (2020). A Connectome of the Adult Drosophila Central Brain. *BioRxiv*, 2020.01.21.911859. <https://doi.org/10.1101/2020.01.21.911859>
- Yan, C., Li, A., Zhang, B., Ding, W., Luo, Q., & Gong, H. (2013). Automated and Accurate Detection of Soma Location and Surface Morphology in Large-Scale 3D Neuron Images. *PLoS One*, 8(4), e62579. <https://doi.org/10.1371/journal.pone.0062579>
- Yin, W., Brittain, D., Borseth, J., Scott, M. E., Williams, D., Perkins, J., Own, C. S., Murfitt, M., Torres, R. M., Kapner, D., Mahalingam, G., Bleckert, A., Castelli, D., Reid, D., Lee, W.-C.A., Graham, B. J., Takeno, M., Bumbarger, D. J., Farrell, C., & da Costa, N. M. (2020). A petascale automated imaging pipeline for mapping neuronal circuits with high-throughput transmission electron microscopy. *Nature Communications*, 11(1), 4949. <https://doi.org/10.1038/s41467-020-18659-3>
- Zeng, H. (2018). Mesoscale Connectomics. *Current Opinion in Neurobiology*, 50, 154–162. <https://doi.org/10.1016/j.conb.2018.03.003>
- Zhou, H., Li, S., Li, A., Huang, Q., & Quan, T. (2021). Gtree: An open-source tool for dense reconstruction of brain-wide neuronal population. *Neuroinformatics*, 19(2), 305–317. <https://doi.org/10.1007/s12021-020-09484-6>
- Zingg, B., Hintiryan, H., Gou, L., Song, M. Y., Bay, M., Bienkowski, M. S., Foster, N. N., Yamashita, S., Bowman, I., Toga, A. W., & Dong, H.-W. (2014). Neural Networks of the Mouse Neocortex. *Cell*, 156(5), 1096–1111. <https://doi.org/10.1016/j.cell.2014.02.023>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.