



PERSPECTIVE

Challenges of Processing and Analyzing Big Data in Mesoscopic Whole-brain Imaging



Anan Li^{1,2,3,a}, Yue Guan^{1,2,b}, Hui Gong^{1,2,3,c}, Qingming Luo^{1,2,*d}

¹ Britton Chance Center for Biomedical Photonics, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China

² MOE Key Laboratory for Biomedical Photonics, School of Engineering Sciences, Huazhong University of Science and Technology, Wuhan 430074, China

³ HUST-Suzhou Institute for Brainsmatics, JITRI Institute for Brainsmatics, Suzhou 215125, China

Received 17 February 2019; revised 15 September 2019; accepted 12 October 2019

Available online 2 December 2019

Handled by Xiangdong Fang

The brain is recognized as the most complex organ on earth. The complex neuronal network in the human brain consists of approximately 86 billion neurons and more than one hundred trillion connections [1]. Even the mouse brain, which is commonly used as a model in neuroscience studies, contains more than 71 million neurons. Despite the great efforts made by the neuroscientists, our knowledge about the brain is still limited and studies on small model animals have only provided preliminary results [2–4]. To improve the situation, brainsmatics studies try to accelerate neuroscience research by providing a standard whole-brain spatial coordinate system and standardized labeling tools, imaging the brain, as well as discovering new neuronal cell types, neuronal circuits, and brain vascular structures.

Neuroscience studies at mesoscopic scale

Spatial brain structure can cross several magnitudes, from synapse, single neuron, and neural circuit to brain nucleus, brain region, and organ [5], as shown in **Figure 1**. Structures at the macroscopic scale include brain regions and white-matter fibers, which are usually larger than 100 μm and can be seen by the naked eye without magnifying optical instruments. Structures that can be viewed only with the help of a microscope, especially an electron microscope, exist at the microscopic scale. In neuroanatomy, structures at microscopic scales are usually made visible with the help of an electron microscope. For example, individual synapses and their vesicles are micro-structures with approximately 100-nm spatial resolution. In addition, there is a mesoscopic scale between the macroscopic and microscopic scales [9]; these mesoscopic structures are typically observable with an optical microscope [10].

Structures at the mesoscopic scale include soma, neurite, and neural circuits, whose diameters are between 100 nm and 100 μm . Neurons and neural circuits at the mesoscopic scale also have macroscopic spatial distribution characteristics [8]. Studies on neurons and neural circuits need to be performed with single-neuron three-dimensional spatial resolution across

* Corresponding author.

E-mail: qluo@mail.hust.edu.cn (Luo Q).

^a ORCID: 0000-0002-5877-4813.

^b ORCID: 0000-0002-3256-3663.

^c ORCID: 0000-0001-5519-6248.

^d ORCID: 0000-0002-6725-9311.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

<https://doi.org/10.1016/j.gpb.2019.10.001>

1672-0229 © 2019 The Authors. Published by Elsevier B.V. and Science Press on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

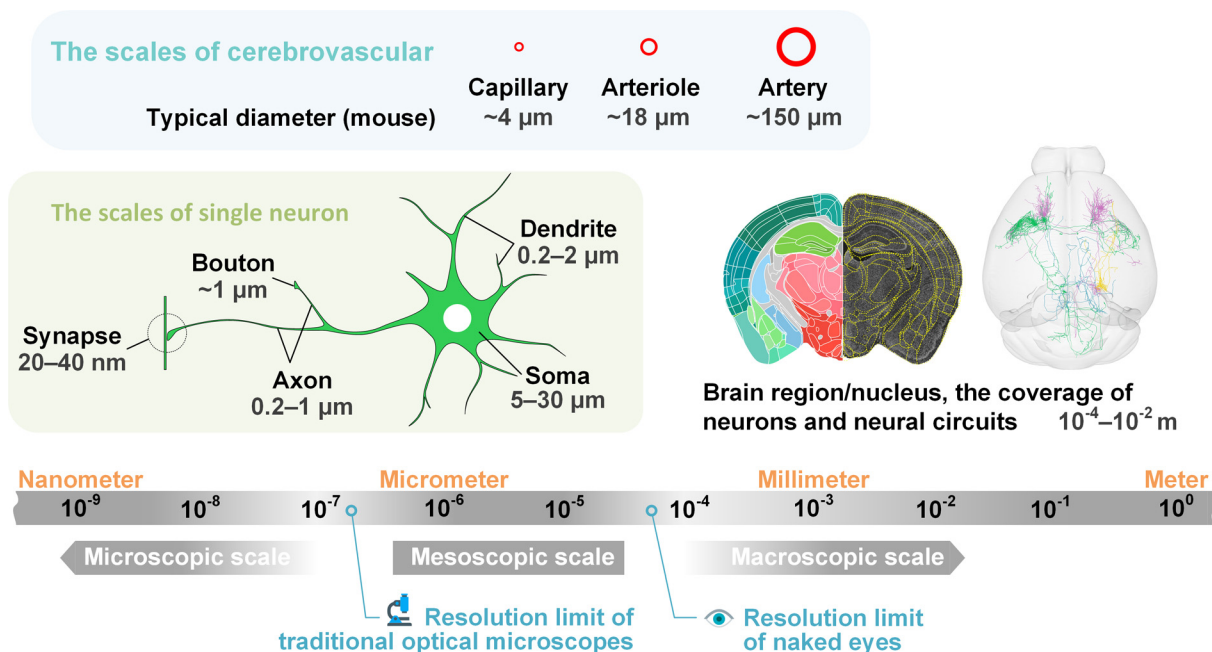


Figure 1 The various scales in neuroscience research

The diameter of a typical capillary blood vessel is about 4 μm and the diameter for arteriole and artery is about 18 μm and 150 μm , respectively. The scale of structures in a single neuron is also at this range: the diameter of a neuron soma is around 5–30 μm ; the diameter of dendrite fiber is around 0.2–2 μm ; the axon fibers have a similar size (0.2–1 μm); however, synapses have much smaller diameter, 20–40 nm. Structures in the cytoarchitecture study, such as brain regions, can span at a range of 0.1–10 mm. The figure is illustrated based on the data reported previously [6–8].

the entire brain [10]; however, reconstructing an entire human brain neuronal network at the mesoscopic scale is similar to taking three-dimensional pictures of a forest with a trillion trees [12], in which we need to see the whole forest as well as each twig and leaf of any single tree. Consequently, constructing whole-brain datasets is a massive undertaking that relies not only on technological advances but also requires massive investments and extensive cooperation [12]. Nonetheless, the data obtained by this approach may help to improve our understanding of the brain. As described previously [13], a comprehensive whole-brain atlas of the cholinergic system was obtained based on a mesoscopic whole-brain dataset, which helps in understanding the role of the cholinergic system in sensory and motor functions in addition to cognitive behaviors.

Brain research projects focus on mesoscopic scale

Driven by the huge community demands, scientists have proposed several global-scale brain research projects [14,15]. The European Human Brain Project (HBP) focuses on collecting and integrating different types of neuroscientific data, reconstructing the brain structure, and simulating brain functions on supercomputers [16,17]. In 2016, HBP released a platform on its official website, providing researchers with brain simulation, visualization, and computational tools. The Brain Research through Advancing Innovative Neurotechnologies (BRAIN) initiative in the United States first develops new methodology and tools, and then attempts to answer scientific questions [18]. The two highest priority goals of the BRAIN initiative are to perform a brain-cell-type census and an atlas

of cross-scale neural circuits [5,11,19]. To march toward these goals, NIH kicked off a pilot program in 2014 and started the BRAIN initiative Cell Census Network (BICCN) in 2017. The task of BICCN is to construct a common whole-mouse brain cell atlas, in which a main focus is to collect information on mesoscopic neuron anatomy [20].

Brain science is also well recognized by the Chinese government and the Chinese research community. “Brain science and cognition” was included in the “Guidelines on National Medium- and Long-term Program for Science and Technology Development” a decade ago. Support is provided through the National Natural Science Foundation of China (NSFC) and the National Basic Research Program of China (973 Program). For example, the key research plan “Neural Circuit Foundation of Emotion and Memory” was supported by NSFC [21]. A Chinese brain project is also in the planning stages, and “brain science and brain-inspired intelligence research” is included as one of the key science and technology innovation programs and projects in the “Thirteenth five-year development plan” and is one of 4 pilot programs of the “science and technique innovation 2030 key program” [22]. Some institutions have been established for Chinese brain projects, such as the Center for Excellence in Brain Science and Intelligence Technology of the Chinese Academy of Sciences [23] and the Chinese Institute for Brain Research in Beijing and Shanghai, respectively [24].

Big data in mesoscopic whole-brain imaging

These projects will generate an enormous quantity of data at an ever-increasing rate; thus, processing and analysis of the

brain big data will likely become a bottleneck [25]. In 2010, we used the Micro-Optical Sectioning Tomography (MOST) system to obtain a Golgi-stained whole-mouse-brain dataset [26]. At $0.3 \mu\text{m} \times 0.3 \mu\text{m} \times 1 \mu\text{m}$ voxel resolution, this dataset contains more than 15,000 coronal images, and the total size of the raw data exceeds 8 TB. The MouseLight project from the Janelia Research Campus uses serial two-photon (STP) tomography to obtain whole-mouse-brain images at a resolution of $0.3 \mu\text{m} \times 0.3 \mu\text{m} \times 1 \mu\text{m}$ voxels; the raw data from each channel is approximately 10 TB [27]. Osten's group from Cold Spring Harbor Laboratory also uses STP tomography to study neural distribution and projection dependent on cell type and to build a brain architecture database [28]. They recently used new oblique light-sheet tomography to obtain a whole-mouse-brain dataset with a raw data size of 11 TB [29].

Human brain volume is approximately 3500 times that of a mouse brain [1]. Thus, mesoscopic scale imaging of the human brain will generate a dataset at the 10 PB scale, as shown in Figure 2. This is equivalent to hundreds of thousands of 4 K movies and equals the storage capacity of the Sunway Taihu-Light supercomputer [30], one of the most powerful supercomputers in existence today. In addition, as the image techniques advance, featuring higher resolutions, more color channels and wider dynamic ranges, systems will generate even larger volumes of data.

Biology labs need a mature and comprehensive solution for managing mesoscopic whole-brain images when studying neuron types and circuit connections. In 2014, *Nature Neuroscience* published a special issue on big data in neuroscience that says, "Neuroscientists have to learn to manage and take advantage of the big waves of data that are being generated." [31]. To distill brain science knowledge from the raw images, the entire processing pipeline typically includes image processing, image registration, information extraction, quantitative

analysis, and visualization, as shown in Figure 3. Consequently, big data bring a range of challenges to data processing, storage, analysis, management, and sharing, because the size of the research target data increases by several folds. A decade of systematic studies on whole-brain imaging from several groups has established a self-contained system to support the applied studies [27,32–37]. Nevertheless, we foresee that there is a long way to go to fully utilize mesoscopic whole-brain images. We recognize the following challenges listed in Table 1 and discuss in the following sections.

The challenge of information extraction and neuron identification

Data segmentation and identification are often recognized as the bottleneck for studies involving big data. The mesoscopic brain-image segmentation and identification target brain regions/nuclei with sizes ranging from hundreds of micrometers to millimeters, neuron soma and micro-vessels with sizes ranging from a few micrometers to 10 micrometers, and boutons with sizes ranging from sub-micrometers to a few micrometers. Axon tracing is one of the most difficult tasks in whole-brain data processing. One reason is that the spatial structures of neurons are complex, the distribution of the neural fibers can be dense, and they can cover large distances but have only a small diameter [38]. Another is that axons and neurites are small structures with weak fluorescent signals that are often difficult to distinguish from the strong background noise. These factors make manual labeling a time-consuming process and degrade the accuracy of automated algorithmic labeling. Thus, manual or semiautomatic axon tracing is still the standard solution in high quality neuron morphology studies. In 2017, *Nature* reported three giant neurons whose fibers cover

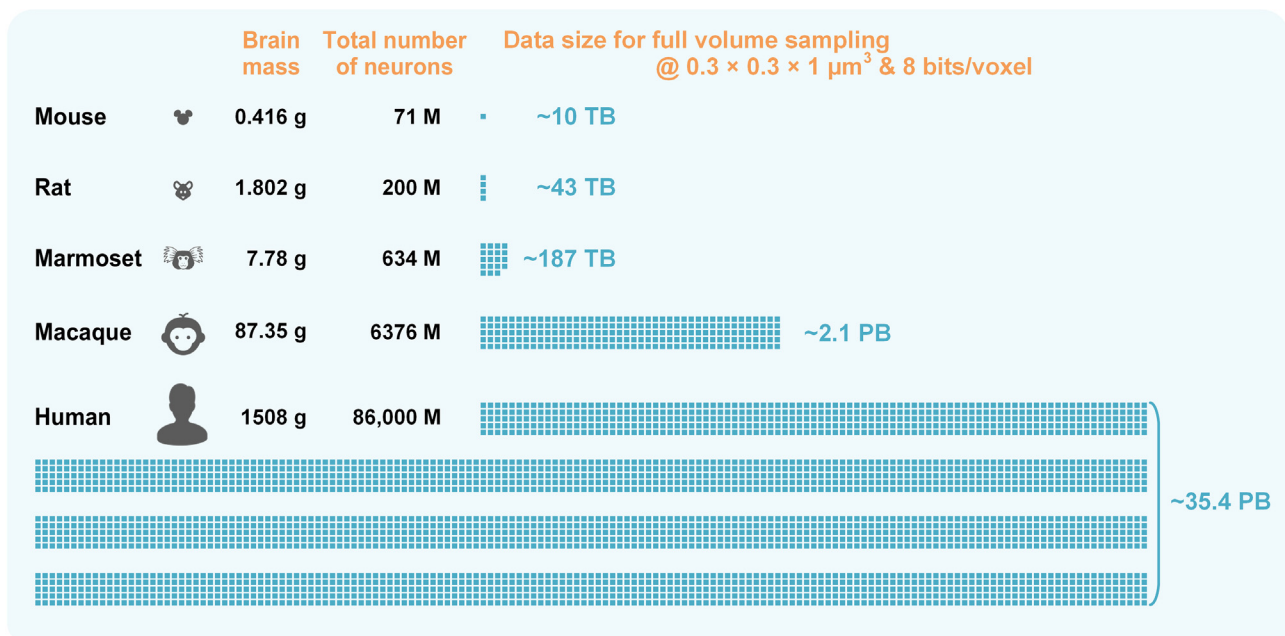


Figure 2 Data size comparison among different species

The brain mass, total number of neurons, and expected data size for full volume sampling for mice, rats, marmosets, macaques, and humans are listed. Each dot in the plot represents 10 TB data. The figure is illustrated based on the data reported previously [1].

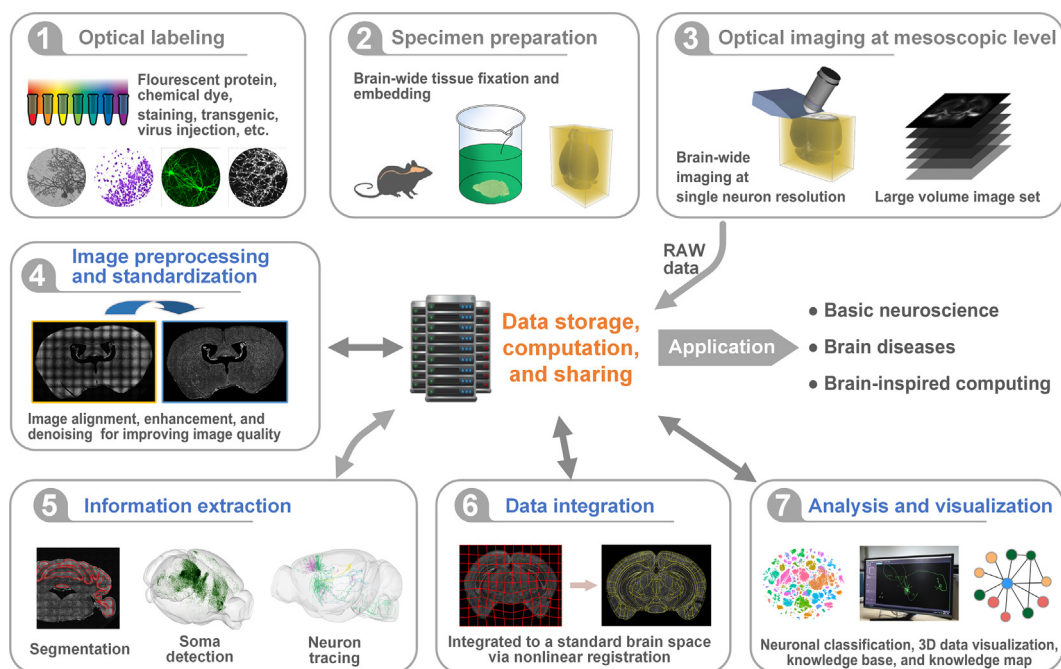


Figure 3 The pipeline of whole-brain data processing and analysis at the mesoscopic scale

The mouse brain is first labeled by fluorescent protein and chemical dye, and then it is fixed and embedded for sectioning. A brain-wide imaging at mesoscopic scale acquires large volume images through fine sectioning. Image preprocessing and standardization of the raw data help to align the images and remove noise. Segmentation, soma detection, and neuron tracing then extract interesting structures from the preprocessed images. Information extracted from different samples and modalities can be integrated through registration to a standard brain space. New knowledge about the brain can be discovered and drawn based on the data from this mesoscopic whole-brain data processing pipeline.

Table 1 Challenges of the whole-brain big data processing and analysis at mesoscopic scale

Tasks	Cause of difficulties	Possible solutions
Information extraction and neuron identification	Complex neuronal morphology and long projection neuronal fibers make neuron identification difficult	<ol style="list-style-type: none"> 1. Accelerating the manual reconstruction process with the help of VR and other techniques 2. Developing deep learning-based techniques with large training datasets
Data integration from multiple omics study	Data collected in multiple omics study don't share the same mode	<ol style="list-style-type: none"> 1. Defining whole brain data point and building a standard whole-brain spatial coordination system around it 2. Registering all modes of data to a standard whole-brain spatial coordination system
Data storage, computation, and sharing	Huge amount of image data bring burden on computation platform	<ol style="list-style-type: none"> 1. Establishing a QC/QA process 2. Developing the distributed computation system to support the whole-brain study at mesoscopic scale
Data and knowledge visualization	Cross-region and cross-scale visualization is tricky	Developing fast data representation techniques to reduce data access and rendering time

Note: VR, virtual reality; QC, quality control; QA, quality assurance.

almost the entire mouse brain, which were all manually reconstructed [39]. Because of the low reliability of automatic reconstruction software, reconstruction may require a day or even a week or more to map a single neuron with highly complex morphology.

Deep learning may help improve the neuron reconstruction issue. However, deep learning currently requires large quantities of labeled data for training, such as ImageNet [40], before

the models are applicable for neuron reconstruction. Another approach attempts to accelerate a semiautomatic process using techniques such as virtual reality [41,42]. Additionally, crowdsourcing may be another approach for solving the human resource problem [43] in data segmentation and neuron fiber identification. Some examples are the project Fold.it for protein study [44] and Openworm for electron microscope data [45].

The challenge of data integration

Data from different samples and different experimental batches need to be integrated—even data obtained through different methodologies and different “omics” can be integrated together for neuroscientific research [46–49]. A whole-brain mesoscopic scale data integration process can employ local, sparse, and cell-type data to build a systematic and comprehensive neuronal network [13,50–52]. However, data from different sources may have nonlinear displacements in spatial coordinates [53,54], which jeopardizes the possible spatial correlation between data samples. Thus, we need a powerful, standardized framework for integrating multi-scale, multi-mode whole-brain big data [15,48,55].

We face the following challenges when building such a framework. First, the data mode often varies, the metadata are scattered, and usually, no standard approach exists for describing the metadata and the data format. The lack of a standard approach introduces unnecessary barriers to the integration and reduces the level of confidence for the integrated data. A second challenge is that no high-resolution whole-brain reference atlas exists for mesoscopic research yet. For example, the Waxholm brain coordinate system is based on magnetic resonance imaging (MRI) data, and its resolution is approximately tens of micrometers [56]. The highest-resolution three-dimensional atlas available is the Allen institute’s Allen mouse common coordinate framework version 3 (CCFv3), whose isotropic resolution is 10 micrometers. The CCFv3 is based on 1675 genetically-engineered mouse brains [34]; the atlas process eliminated individual differences by averaging the brain tissue’s auto-fluorescence signals [57]. Finally, no registration method currently exists that can accurately map mesoscopic whole-brain datasets to a standard brain space. While some existing registration methods work well on small regions of the mesoscopic scale data, no current methods are able to work on 10-TB data scales to our knowledge [58]. An ideal anatomical datum is useful because it can be used to define a standard brain space coordinate system despite sample and development varieties. However, the existing conservative anatomical data are small and subject to manual biases.

The challenge of data storage, computation, and sharing

The quality of data used in neuroscience research can be categorized. The BICCN project classifies the data into five categories: raw, quality control/quality assurance (QA/QC), linked, featured, and integrated (<https://biccn.org>). The mesoscopic whole-brain data size in the raw, QA/QC, and linked categories is quite large, imposing a huge burden on data storage, computation, and sharing. One important approach for confronting this whole-brain big data challenge is to reduce the data size when not all the raw data are required. For example, we can try to capture high signal-to-noise ratio raw data in the imaging stage, skip imaging in uninterested areas, as well as downsample, denoise, and compress the data. Data compression in particular may be a valid option because considerable redundant information exists in raw three-dimensional whole-brain image data [59]. The data compression techniques used in video compression [27] and compressed sensing [60]

may also be applicable for brain imaging data. Due to the large volume of mesoscopic-scale whole-brain data, even small improvements in data compression can have large effects in relieving the requirements for data storage, computation, and transmission.

Another concern with mesoscopic whole-brain data processing is in big data QC, as expanding studies involve international cooperation and data integration efforts. The quality standards for both raw and derived data are prerequisites for productive data exchange and cooperation between research teams. Scientists need easy-to-use processing, analysis, and QA/QC tools to adhere to and benefit from these quality standards [61–63]. The International Neuroinformatics Coordinating Facility (INCF) is a non-profit science organization established in 2005 that acts as the most important facilitator for standards, tools, and sharing of big data in neuroscience between the large international brain initiatives [64]. Individual scientific software applications can register with INCF Central to identify their file formats, transformations, standard query formats, and other essential metadata [48].

In addition, an increasing number of desktop applications will migrate to cloud platforms, especially those involving biomedical big data at PB/EB scales (1024 PB equals 1 EB) [65,66]. Over a cloud platform, users can use client applications to send commands to the cloud platform; then, a cloud server executes the requested data manipulations, computations, and visualizations. Centralizing the data in this way obviates the need to transport the big data; thus, this mode reduces the transmission burden and fosters unified data formats and quality standards. However, building and running big data cloud platforms requires a huge upfront investment, coordination with other existing computation facilities, as well as community financial and management support.

The challenge of data and knowledge visualization

Eventually, mesoscopic whole-brain image data are refined into vectorized graphic data via preprocessing, segmentation, labeling, and integration processes [67]. These graphic data contain precise spatial positioning information, forming the basic data needed for further studies, such as neuron cell typing and connections between brain regions [68,69]. Visualization is a necessary aspect of these studies. Visualizations present complex data using graphics tools, including symbols, colors, and textures. These graphics tools can deliver information effectively and promote data salience. Using visualization tools makes interpreting, manipulating, and processing data simpler and allows patterns, characteristics, correlation, and anomalies to be recognized or constructed. Although visualization has been widely applied in biomedical studies, the large volume, cross-magnitude, high-dimensional, and complex mesoscopic whole-brain data bring new challenges to the existing visualization techniques. For example, both main memory and graphics memory become bottlenecks when three-dimensional data at 10 TB or PB scales need to be rendered [33,70,71]. Visualization applications should support rendering from 10 K to millions of neurons, allowing both mesoscopic neuron morphological detail and macroscopic brain region connections to be viewed. Visualization should also support various types of data manipulation and presentation, including oblique reslicing [29], immersive display [72], neuron recon-

struction analysis [35], brain connection atlas presentation [34], and visualizations distributed to clients by clouds [73].

Promote brain-inspired artificial intelligence

Artificial intelligence (AI) has achieved remarkable progresses after more than 6 decades of development. The defeat of the human Go champion by Google's AlphaGo in 2016 ushered in a new wave of intelligent techniques [74]. However, developments in AI are limited, have weak intelligence, and models cannot be applied in scenarios other than the one the model was designed to address. One reason may be that AI studies have not yet fully learned how brain mechanisms work, and many are not brain-inspired [75]. The Blue Brain Project from Ecole polytechnique fédérale de Lausanne (EPFL) tries to simulate mouse brain cortex activity based on neuron morphology data, electrophysiology data, and so on [76,77]. The BRAIN Initiative launched the Machine Intelligence from Cortical Networks (MICrONS) program to reconstruct neural circuits in a 1-cubic millimeter brain volume, simulate cortical function, and develop next-generation machine intelligence systems [78]. The forthcoming China Brain Initiative also prioritizes brain-inspired AI over other lower priority long-term subjects [75]. Neurons are the basic information processing unit in the brain, and neural circuits are the basic structure involved in brain function. Thus, the fine spatial–temporal structure and functional data collected via mesoscopic studies can provide clues in understanding how brains function and promote brain-inspired AI efforts [79].

The research and application of mesoscopic whole-brain image data is just beginning, and many challenges and difficulties lie ahead. We hope to see increasing cooperation between scientists and engineers for the various brain projects in the future.

Competing interests

The authors have declared that no competing interests exist.

Acknowledgments

This study is supported by Science Fund for Creative Research Group of the National Natural Science Foundation of China (Grant No.61721092), and the National Natural Science Foundation of China (Grant No. 61890954).

References

- [1] Herculano-Houzel S. The human brain in numbers: a linearly scaled-up primate brain. *Front Hum Neurosci* 2009;3:31.
- [2] Hildebrand DGC, Cicconet M, Torres RM, Choi W, Quan TM, Moon J, et al. Whole-brain serial-section electron microscopy in larval zebrafish. *Nature* 2017;545:345.
- [3] Gao R, Asano SM, Upadhyayula S, Pisarev I, Milkie DE, Liu TL, et al. Cortical column and whole-brain imaging with molecular contrast and nanoscale resolution. *Science* 2019;363:eaau8302.
- [4] Erö C, Gewaltig M-O, Keller D, Markram H. A cell atlas for the mouse brain. *Front Neuroinform* 2018;12:84.
- [5] Lichtman JW, Denk W. The Big and the small: challenges of imaging the brain's circuits. *Science* 2011;334:618–23.
- [6] Muller B, Lang S, Dominietto M, Rudin M, Schulz G, Deyhle H, et al. High-resolution tomographic imaging of microvessels. *Proc SPIE* 7078, Developments in X-ray tomography VI 2008:70780B.
- [7] Snyder EY, Yoon C, Flax JD, Macklis JD. Multipotent neural precursors can differentiate toward replacement of neurons undergoing targeted apoptotic degeneration in adult mouse neocortex. *Proc Natl Acad Sci U S A* 1997;94:11663–8.
- [8] Stuart G, Spruston N, Häusser M. *Dendrites*. Oxford: Oxford University Press; 2016.
- [9] Bohland JW, Wu C, Barbas H, Bokil H, Bota M, Breiter HC, et al. A proposal for a coordinated effort for the determination of brainwide neuroanatomical connectivity in model organisms at a mesoscopic scale. *PLoS Comput Biol* 2009;5:e1000334.
- [10] Osten P, Margrie TW. Mapping brain circuitry with a light microscope. *Nat Methods* 2013;10:515–23.
- [11] Mitra P. The Circuit architecture of whole brains at the mesoscopic scale. *Neuron* 2014;83:1273–83.
- [12] DeFelipe J. Neuroanatomy and global neuroscience. *Neuron* 2017;95:14–8.
- [13] Li X, Yu B, Sun Q, Zhang Y, Ren M, Zhang X, et al. Generation of a whole-brain atlas for the cholinergic system and mesoscopic projectome analysis of basal forebrain cholinergic neurons. *Proc Natl Acad Sci U S A* 2018;115:415–20.
- [14] Grillner S, Ip N, Koch C, Koroshetz W, Okano H, Polachek M, et al. Worldwide initiatives to advance brain research. *Nat Neurosci* 2016;19:1118.
- [15] Huang ZJ, Luo L. It takes the world to understand the brain. *Science* 2015;350:42–4.
- [16] Amunts K, Ebell C, Muller J, Telefont M, Knoll A, Lippert T. The human brain project: creating a european research infrastructure to decode the human brain. *Neuron* 2016;92:574–81.
- [17] Koslow SH, Huerta MF. *Neuroinformatics: an overview of the human brain project*. Psychology Press; 2013.
- [18] Insel TR, Landis SC, Collins FS. The NIH BRAIN Initiative. *Science* 2013;340:687–8.
- [19] Zeng H. Mesoscale connectomics. *Curr Opin Neurobiol* 2018;50:154–62.
- [20] Ecker JR, Geschwind DH, Kriegstein AR, Ngai J, Osten P, Polioudakis D, et al. The BRAIN Initiative cell census consortium: lessons learned toward generating a comprehensive brain cell atlas. *Neuron* 2017;96:542–57.
- [21] Zhu Y, Cao H, Dong E. Grand research plan for neural circuits of emotion and memory — current status of neural circuit studies in China. *Neurosci Bull* 2013;29:121–4.
- [22] Poo M, Du J, Ip Nancy Y, Xiong Z, Xu B, Tan T. China brain project: basic neuroscience, brain diseases, and brain-inspired computing. *Neuron* 2016;92:591–6.
- [23] Qiu J. Research and development of artificial intelligence in China. *Natl Sci Rev* 2016;3:538–41.
- [24] Cyranoski D. Beijing launches pioneering brain-science centre. *Nature* 2018;556:157–8.
- [25] Frégnac Y. Big data and the industrialization of neuroscience: a safe roadmap for understanding the brain? *Science* 2017;358:470–7.
- [26] Li A, Gong H, Zhang B, Wang Q, Yan C, Wu J, et al. Micro-optical sectioning tomography to obtain a high-resolution atlas of the mouse brain. *Science* 2010;330:1404–8.
- [27] Economo MN, Clack NG, Lavis LD, Gerfen CR, Svoboda K, Myers EW, et al. A platform for brain-wide imaging and reconstruction of individual neurons. *eLife* 2016;5:e10566.
- [28] Kim Y, Yang GR, Pradhan K, Venkataraju KU, Bota M, García del Molino LC, et al. Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell* 2017;171:456–69.e22.
- [29] Narasimhan A, Umadevi Venkataraju K, Mizrachi J, Albeanu DF, Osten P. Oblique light-sheet tomography: fast and high resolution volumetric imaging of mouse brains. *bioRxiv* 2017;132423.

- [30] Fu H, Liao J, Yang J, Wang L, Song Z, Huang X, et al. The sunway taihulight supercomputer: system and applications. *Sci China Inform Sci* 2016;59:072001.
- [31] Focus on big data. *Nat Neurosci* 2014;17:1429.
- [32] Amat F, Höckendorf B, Wan Y, Lemon WC, McDole K, Keller PJ. Efficient processing and analysis of large-scale light-sheet microscopy data. *Nat Protoc* 2015;10:1679.
- [33] Li Y, Gong H, Yang X, Yuan J, Jiang T, Li X, et al. TDat: an efficient platform for processing petabyte-scale whole-brain volumetric images. *Front Neural Circuits* 2017;11:55.
- [34] Kuan L, Li Y, Lau C, Feng D, Bernard A, Sunkin SM, et al. Neuroinformatics of the allen mouse brain connectivity atlas. *Methods* 2015;73:4–17.
- [35] Shillcock JC, Hawrylycz M, Hill S, Peng H. Reconstructing the brain: from image stacks to neuron synthesis. *Brain Inform* 2016;3:205–9.
- [36] Freeman J. Open source tools for large-scale neuroscience. *Curr Opin Neurobiol* 2015;32:156–63.
- [37] Amunts K, Lepage C, Borgeat L, Mohlberg H, Dickscheid T, Rousseau MÉ, et al. BigBrain: an ultrahigh-resolution 3D human brain model. *Science* 2013;340:1472–5.
- [38] Gerfen CR, Economo MN, Chandrashekar J. Long distance projections of cortical pyramidal neurons. *J Neurosci Res* 2018;96:1467–75.
- [39] Reardon S. A giant neuron found wrapped around entire mouse brain. *Nature* 2017;543:14–5.
- [40] Deng J, Dong W, Socher R, Li LJ, Li K, Li FF. ImageNet: a large-scale hierarchical image database. *IEEE Conf Comput Vis Pattern Recognit* 2009;1–4:248–55.
- [41] Boorboor S, Jadhav S, Ananth M, Talmage D, Role L, Kaufman A. Visualization of neuronal structures in wide-field microscopy brain images. *IEEE Trans Vis Comput Graph* 2019;25:1018–28.
- [42] Usher W, Klacansky P, Federer F, Bremer P, Knoll A, Yarch J, et al. A virtual reality visualization tool for neuron tracing. *IEEE Trans Vis Comput Graph* 2018;24:994–1003.
- [43] Roskams J, Popović Z. Power to the people: addressing big data challenges in neuroscience by creating a new cadre of citizen neuroscientists. *Neuron* 2016;92:658–64.
- [44] Khatib F, Cooper S, Tyka MD, Xu K, Makedon I, Popović Z, et al. Algorithm discovery by protein folding game players. *Proc Natl Acad Sci U S A* 2011;108:18949–53.
- [45] Szigeti B, Gleeson P, Vella M, Khayrulin S, Palyanov A, Hokanson J, et al. OpenWorm: an open-science approach to modeling *Caenorhabditis elegans*. *Front Comput Neurosci* 2014;8:137.
- [46] Bjerke IE, Øvsthus M, Papp EA, Yates SC, Silvestri L, Fiorilli J, et al. Data integration through brain atlas: human brain project tools and strategies. *Eur Psychiat* 2018;50:70–6.
- [47] Gupta A, Ludäscher B, Martone ME. Knowledge-based integration of neuroscience data sources.. 12th Int Conf Sci Stat Database Manag 2000:39–52.
- [48] Hawrylycz M, Baldock RA, Burger A, Hashikawa T, Johnson GA, Martone M, et al. Digital atlas and standardization in the mouse brain. *PLoS Comput Biol* 2011;7:e1001065.
- [49] Boline J, Lee EF, Toga A. Digital atlases as a framework for data sharing. *Front Neurosci* 2008;2:12.
- [50] Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, et al. A mesoscale connectome of the mouse brain. *Nature* 2014;508:207.
- [51] Peng J, Long B, Yuan J, Peng X, Ni H, Li X, et al. A quantitative analysis of the distribution of CRH neurons in whole mouse brain. *Front Neuroanat* 2017;11:63.
- [52] Zhang C, Yan C, Ren M, Li A, Quan T, Gong H, et al. A platform for stereological quantitative analysis of the brain-wide distribution of type-specific neurons. *Sci Rep* 2017;7:14334.
- [53] Zitová B, Flusser J. Image registration methods: a survey. *Image Vision Comput* 2003;21:977–1000.
- [54] Kakadiaris IA, Bello M, Arunachalam S, Kang W, Ju T, Warren J, et al. Landmark-driven, atlas-based segmentation of mouse brain tissue images containing gene expression data. *Med Image Comput Assist Interv* 2004:192–9.
- [55] Landhuis E. Neuroscience: big brain, big data. *Nature* 2017;541:559.
- [56] Johnson GA, Badea A, Brandenburg J, Cofer G, Fubara B, Liu S, et al. Waxholm space: an image-based reference for coordinating mouse brain research. *NeuroImage* 2010;53:365–72.
- [57] Allen Institute for Brain Science. Allen mouse common coordinate framework: allen mouse common coordinate framework and reference atlas. Technical White Paper 2017.
- [58] Ni H, Tan C, Feng Z, Chen S, Zhang Z, Li W, et al. A robust image registration interface for large volume brain atlas. *bioRxiv* 2018;377044.
- [59] Bilgin A, Zweig G, Marcellin MW. Three-dimensional image compression with integer wavelet transforms. *Appl Optics* 2000;39:1799–814.
- [60] Donoho DL. Compressed sensing. *IEEE Trans Inform Theory* 2006;52:1289–306.
- [61] Ferguson AR, Nielson JL, Cragin MH, Bandrowski AE, Martone ME. Big data from small data: data-sharing in the ‘long tail’ of neuroscience. *Nat Neurosci* 2014;17:1442.
- [62] Martone ME, Gupta A, Ellisman MH. E-neuroscience: challenges and triumphs in integrating distributed data from molecules to brains. *Nat Neurosci* 2004;7:467–72.
- [63] Poldrack RA, Gorgolewski KJ. Making big data open: data sharing in neuroimaging. *Nat Neurosci* 2014;17:1510–7.
- [64] Bjaalie JG, Grillner S. Global neuroinformatics: the international neuroinformatics coordinating facility. *J Neurosci* 2007;27:3613–5.
- [65] Grossman RL, White KP. A vision for a biomedical cloud. *J Intern Med* 2012;271:122–30.
- [66] Rosenthal A, Mork P, Li MH, Stanford J, Koester D, Reynolds P. Cloud computing: a new business paradigm for biomedical information sharing. *J Biomed Inform* 2010;43:342–53.
- [67] Parekh R, Ascoli Giorgio A. Neuronal morphology goes digital: a research hub for cellular and system neuroscience. *Neuron* 2013;77:1017–38.
- [68] Han Y, Kechschull JM, Campbell RAA, Cowan D, Imhof F, Zador AM, et al. The logic of single-cell projections from visual cortex. *Nature* 2018;556:51.
- [69] Lin R, Wang R, Yuan J, Feng Q, Zhou Y, Zeng S, et al. Cell-type-specific and projection-specific brain-wide reconstruction of single neurons. *Nat Methods* 2018;15:1033–6.
- [70] Bria A, Iannello G, Onofri L, Peng HC. TeraFly: real-time three-dimensional visualization and annotation of terabytes of multi-dimensional volumetric images. *Nat Methods* 2016;13:192–4.
- [71] Pietzsch T, Saalfeld S, Preibisch S, Tomancak P. BigDataViewer: visualization and processing for large image data sets. *Nat Methods* 2015;12:481–3.
- [72] Matusik W, Pfister H. 3D TV: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *Acm T Graphic* 2004;23:814–24.
- [73] Dutra M, Rodrigues PSS, Giraldo GA, Schulze B. Distributed visualization using VTK in grid environments. 7th IEEE Int Symposium Cluster Comput Grid 2007:381–8.
- [74] Wang F, Zhang JJ, Zheng X, Wang X, Yuan Y, Dai X, et al. Where does AlphaGo go: from church-turing thesis to AlphaGo thesis and beyond. *IEEE-CAA J Automatic* 2016;3:113–20.
- [75] Poo M. Towards brain-inspired artificial intelligence. *Natl Sci Rev* 2018;5:785.
- [76] Markram H, Muller E, Ramaswamy S, Reimann Michael W, Abdellah M, Sanchez Carlos A, et al. Reconstruction and simulation of neocortical microcircuitry. *Cell* 2015;163:456–92.
- [77] Markram H. The blue brain project. *Nat Rev Neurosci* 2006;7:153.
- [78] Cepelewicz J. The US government launches a \$100-million “apollo project of the brain”. *Sci Am* 2016, March 08. Available at <https://www.scientificamerican.com/article/the-u-s-government-launches-a-100-million-apollo-project-of-the-brain/>.
- [79] Luo Q. Brainsmatics—bridging the brain science and brain-inspired artificial intelligence. *Sci Sin Vitae* 2017;47:1015–24.