

# Review on data analysis methods for mesoscale neural imaging *in vivo*

Yeyi Cai, Jiamin Wu<sup>✉,\*</sup> and Qionghai Dai\*

Tsinghua University, Department of Automation, Beijing, China

## Abstract

**Significance:** Mesoscale neural imaging *in vivo* has gained extreme popularity in neuroscience for its capacity of recording large-scale neurons in action. Optical imaging with single-cell resolution and millimeter-level field of view *in vivo* has been providing an accumulated database of neuron-behavior correspondence. Meanwhile, optical detection of neuron signals is easily contaminated by noises, background, crosstalk, and motion artifacts, while neural-level signal processing and network-level coordinate are extremely complicated, leading to laborious and challenging signal processing demands. The existing data analysis procedure remains unstandardized, which could be daunting to neophytes or neuroscientists without computational background.

**Aim:** We hope to provide a general data analysis pipeline of mesoscale neural imaging shared between imaging modalities and systems.

**Approach:** We divide the pipeline into two main stages. The first stage focuses on extracting high-fidelity neural responses at single-cell level from raw images, including motion registration, image denoising, neuron segmentation, and signal extraction. The second stage focuses on data mining, including neural functional mapping, clustering, and brain-wide network deduction.

**Results:** Here, we introduce the general pipeline of processing the mesoscale neural images. We explain the principles of these procedures and compare different approaches and their application scopes with detailed discussions about the shortcomings and remaining challenges.

**Conclusions:** There are great challenges and opportunities brought by the large-scale mesoscale data, such as the balance between fidelity and efficiency, increasing computational load, and neural network interpretability. We believe that global circuits on single-neuron level will be more extensively explored in the future.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.NPh.9.4.041407](https://doi.org/10.1117/1.NPh.9.4.041407)]

**Keywords:** data analysis; mesoscale neural imaging; pipeline.

Paper 21077SSVR received Dec. 30, 2021; accepted for publication Mar. 23, 2022; published online Apr. 15, 2022.

## 1 Introduction

Recording neural activities *in vivo* with optical systems and genetically encoded fluorescence indicators provides an observation window for neural scientists to understand the signal processing procedure of individual neurons and the circuitry of neural network in action. Compared with electrophysiological methods, optical imaging *in vivo* is typically less invasive and could record several brain areas up to millimeter-level field of view (FOV) at cellular resolution.<sup>1,2</sup> Animal surgeries, such as cranial window, thinned skull, or crystal skull<sup>3,4</sup> provide the optical imaging window for one-photon imaging to achieve single-cell resolution in the superficial cortex, such as layer 2/3. Optical neural imaging has thus been used to investigate neural structure change,<sup>5,6</sup> brain state alteration,<sup>7</sup> and information flow while the animal performs a specific task. Profound

---

\*Address all correspondence to Qionghai Dai, [qhdai@tsinghua.edu.cn](mailto:qhdai@tsinghua.edu.cn); Jiamin Wu, [wujiamin@tsinghua.edu.cn](mailto:wujiamin@tsinghua.edu.cn)

discovery has been made by *in vivo* neural imaging on crucial neuroscience issues, including perceptual input processing,<sup>8,9</sup> motion,<sup>10,11</sup> learning,<sup>12,13</sup> memory,<sup>14,15</sup> and decision making.<sup>16</sup>

To date, multiple imaging modalities are capable of mesoscale recording with single-cell resolution and millimeter-level FOV, such as single-photon widefield microscopy,<sup>3,17</sup> multi-photon microscopy,<sup>18,19</sup> light-field microscopy,<sup>20,21</sup> and light-sheet microscopy.<sup>22</sup> A recent review<sup>2</sup> has described the mesoscale imaging techniques and related animal models in detail, which also points out its increasing importance in neuroscience. While the challenges and opportunities brought by the large-scale mesoscale data have been gradually realized by the community,<sup>2</sup> there is still not a comprehensive review about the existing mesoscale analysis methods, remaining problems, and potential future directions. Here, we intended to provide a general data analysis pipeline for cellular level mesoscale neural imaging without focusing on any specific imaging modality. We emphasize the commonalities between different image systems. For instance, the data analysis always starts with neural images from detectors and ends by useful information extracted from calcium signals. Some imaging modalities have their unique preprocess algorithm before acquiring the neural image, such as slice stitching in light-sheet microscopy and volume reconstruction in light-field microscopy. These are beyond the scope of this work. However, we do illustrate the specific priors in data analysis processing, which can be considered based on different imaging modalities, during our detailed descriptions of each processing step. For readers who might be not familiar with microscopy systems, we briefly review the image modalities used for *in vivo* neural dynamic imaging in the following paragraphs.

In the past decades, benefitting from the rapid development of both microscopic systems<sup>17,23–25</sup> and fluorescent indicators,<sup>26,27</sup> *in vivo* neural imaging has been extending its capability in faster sampling speed, higher resolution, larger FOV, and lower phototoxicity.<sup>28</sup> The simple wide-field microscopy could cover several adjacent cortex areas,<sup>29</sup> but it does not typically achieve cellular resolution because of scattering and aberration. Several recent works have shown capabilities to extract single-cell information in wide-field data even with strong background fluorescence by matrix factorization and deep learning.<sup>30,31</sup> In addition, animal models and new fluorescence indicators with specific labeling strategies can greatly reduce the background fluorescence in normal wide-field microscopes, facilitating single-cell resolution neural recoding with simple systems, e.g., layer-specific labeling,<sup>3</sup> and soma-targeted sensors.<sup>32,33</sup> In addition, with special optical designs, several works have been done to further increase the resolution or depth of field with better fidelity to retrieve the single-cell neural traces. For instance, in the RUSH system,<sup>17</sup> a  $5 \times 7$  camera array was tiled to cover centimeter-level FOV and reach  $0.8\text{-}\mu\text{m}$  resolution with dense sampling density and layer-specific neuron labeling; the COSMOS microscope<sup>4</sup> uses multifocal optical sampling to record in-focus projection of  $1\text{ cm} \times 1\text{ cm} \times 1.3\text{ mm}$  volume at near cellular resolution (1–15 neurons/unit). To examine the on-focus slice only, confocal microscopy<sup>34</sup> and light sheet<sup>35</sup> microscopy were designed by either blocking out the out-of-focus light or illuminating a thin slice of the tissue from the side. These techniques have enabled optically sectioning of the brain tissue with three-dimensional (3D) resolving power through scanning strategy. Nevertheless, the resolution of single-photon microscopy degrades tremendously with the increase of penetration depth,<sup>36</sup> thus it was only useful in detecting neurons in shallow cortex layers. Multiphoton microscopy (MPM), on the other hand, holds the advantage in penetration depth<sup>37</sup> and low photodamage. Multiple photons with lower energy cooperate to excite the fluorophore with a nonlinear absorption rate to light intensity. Hence, MPM was widely used to image deep mouse brain.<sup>38,39</sup> Gradient index microlenses further extended the imaging depth to even deeper brain nuclear by affecting the optical path of the exited fluorescence.<sup>40</sup>

There have been emerging computational techniques for high-throughput 3D volumetric imaging.<sup>41</sup> In two-photon microscopy, 3D imaging is accomplished by quickly scanning the sample with single-dot or slice excitation which could be either sequentially or randomly.<sup>42</sup> Yet the sampling frequency is essentially limited by the control frequency of the mechanical actuator and the inertia of the optical system. An alternative approach was through multiplexing. A multi-focus microscope acquires multiple depth information by multi-focusing optical path<sup>19,43–45</sup> or point-spread function engineering.<sup>18,46</sup> Light-field microscopy captures the 3D information efficiently in a tomographic manner with extended depth of field along different

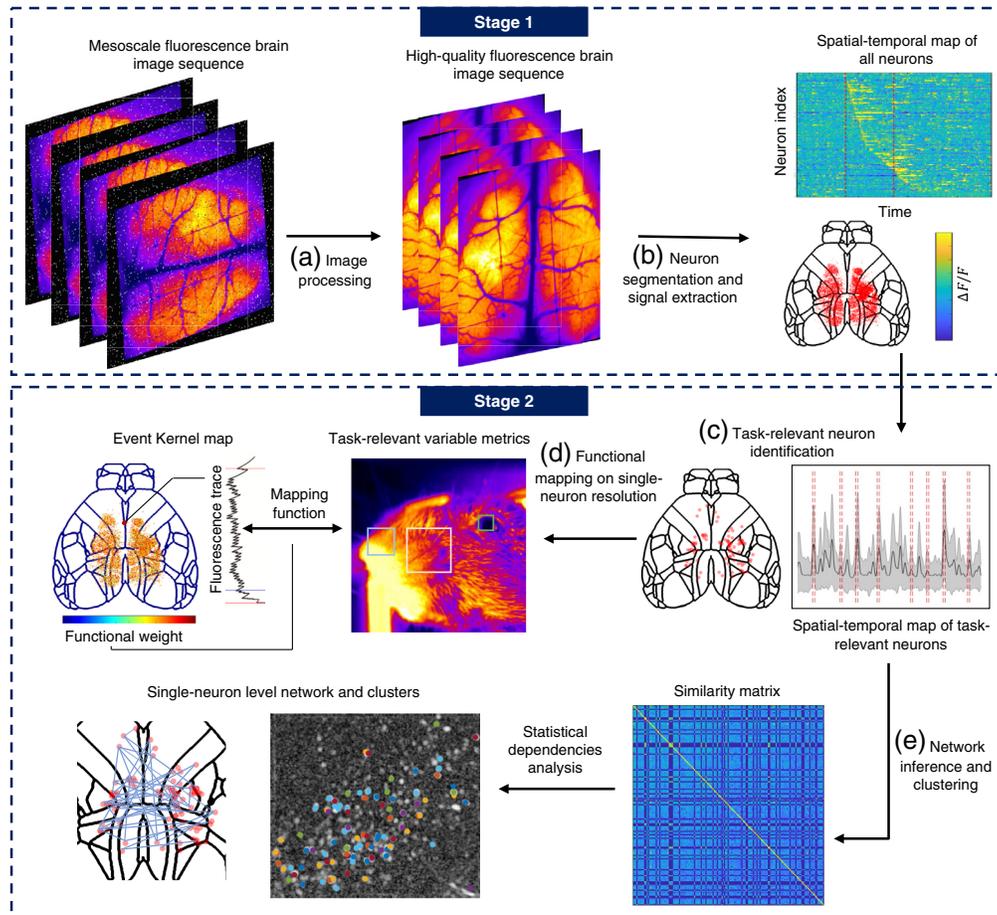
angles.<sup>47–49</sup> While scanning light-field microscopy significantly increases the spatial resolution in multi-cellular organisms,<sup>20</sup> confocal light-field microscopy,<sup>21</sup> or computational optical sectioning<sup>50</sup> further increase its signal-to-background ratio in brain tissue. These techniques have enabled parallel volumetric imaging, capturing over thousands of neurons at cellular level with dozens of volumes per second.

Apart from benchtop microscopy systems where animals are head-fixed, head-mounted miniature microscopies would allow animals to freely move in experimental environments. Miniature microscopes could facilitate studies that are better performed in unrestrained subjects, such as spatial navigation, social behavior, and reward-seeking. To minimize the weight and size of the system, light-emitting diode, image acquisition chips, and miniaturized lenses are commonly used in miniature microscopy. Progress has been made in miniature systems with millimeter-level FOV and near cellular resolution.<sup>51,52</sup>

Unlike electrophysiology detection, optical detection relies on the photon transmission of the genetically encoded indicators and the optical sensor. The optics and electric conversion occurred twice during the imaging process, once by optical bio-indicator, once by the camera sensor. The indirectness of the signal detection results in potential signal corruption and reconstruction indispensability. In the meantime, mesoscale neural imaging usually features with an extremely large data throughput across multiple scales. These barriers of signal recovery have strengthened the essentiality of efficient and accurate computational approaches to extract high-fidelity neural activities from large-scale raw images captured by mesoscale imaging systems.

Here, we review recent data analysis methods for mesoscale intravital neural imaging along a general data-processing pipeline divided into two main stages (Fig. 1). Stage 1 includes several image processing procedures and outputs the compressed spatial-temporal single-neuron traces.<sup>53,54</sup> Stage 2 includes various data-mining methods to interpret mesoscale neural signals both at the cellular level and network level. First, the video captured by the camera sensor should be registered to a template. Image sequences can be motion-blurred because of the heart-beating, breathing, or moving gestures of the animal. To identify specific neurons across a long term, all frames should be registered to a reference position. The second step is image denoising, the method of which depends on different signal-noise ratios (SNRs) and imaging modalities. Lower laser power is always preferred with *in vivo* experiments due to phototoxicity, in which case Poisson noise usually dominates over the readout noise and dark noise with high-speed high-sensitivity detectors. Under low-light conditions, computational denoising methods become indispensable because the noise can easily corrupt the down-stream analysis and interfere the interpretation of the neural activities. Calcium fluorescence signals might also be interfered by hemodynamic, which should be corrected before signal extraction. After this, the 4D/3D stack [3D/2D spatially and one-dimensional (1D) temporally] of the brain tissue is ready for neuron signal extraction. The goal of signal extraction is the demix of spatial-temporal information embedded in the fluorescence image. Neurons need to be spatially segmented from the brain tissue background, and their temporal traces are extracted from the temporal sequence of the image stack. These two steps could be performed sequentially by first determining the position or footprint of the neurons and averaging the relevant pixels for temporal traces, or parallelly by treating the spatial-temporal dimension as equivalent dimensions and utilizing the low-rank prior to perform a tensor-factorization. After the signal extraction step, the data size should be reduced to several megabytes (MBs) and could be represented as two 2D matrixes containing the temporal trace of neurons and their spatial footprints. The following analysis of the neuron trace could be diverse depending on the research problem. Generally, there are two levels of analysis. The first level is to investigate the single-neuron property, such as their tuning curves and the post-stimulus time histogram. The reaction patterns of these neurons may be used to analyze their relationship to certain stimuli or behavior. The second level of analysis extends the scale to local or global circuits formed by single neurons. These kinds of studies aim to infer the mesoscale functional network connection between the neurons to reconstruct the calculating strategy the neural circuit uses to accomplish specific signal processing missions.

In the following sections, we will review recent data analysis methods in each step. And we classify and evaluate different approaches in algorithm feasibility, scope of application, and precisions. The first stage which aims at a precise and efficient extraction of neuron traces is



**Fig. 1** General pipeline of analyzing mesoscale fluorescence functional neural images. The whole pipeline is summarized into two main stages. The first stage targeted at spatial-temporal demixing of neural signals and the second stage targeted at data mining. Each stage contains a sequence of processes, which is framed up with a dashed-line box. (a) Image preprocesses includes three steps: motion registration, denoising, and hemodynamic correction. (b) After calcium signal extraction, the raw video sequence is decomposed into the spatial information of each neuron and their temporal fluorescence signals. (c) Task-relevant neurons only make up a small proportion of total neurons, thus should be recognized through statistical tests. (d) Linear and non-linear regressions map the test-relevant variables such as performance accuracy, gesture and choice into single-neuron traces, thus revealing the functional role of each neuron. (e) Similarity matrix based on correlation, cross entropy and causality could be used to analyze statistical dependencies between neurons and induce the neuron cluster community and neural network.

introduced in Sec. 2; in Sec. 3, we introduce the second stage which explores the underlying neural property and functional circuits through these mesoscale imaging data. Finally, in Sec. 4 we discuss the prospects and remaining challenges for mesoscale neural signal analysis.

## 2 Image Processing and Signal Extraction

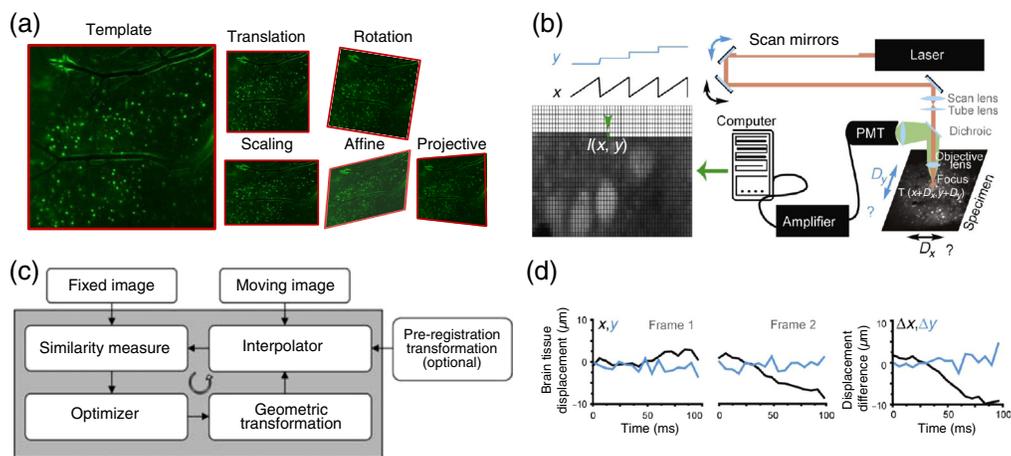
### 2.1 Image Preprocess: Motion Registration, Denoise, and Hemodynamic Correction

Calcium imaging is often accompanied by motion artifacts, even if the animal was head-fixed and anesthetized. In head-fixed experiments, the non-rigid warping of the brain tissue could be caused by heart-beating, breathing, or the shrinking of the tissue from exposure to the laser. Moreover, with freely moving animals, the motion artifact becomes even more severe, because

the posture changes induce relative movement of the head and objective lens. In long-term observation experiments, which expand several hours or days, as well as experiments involving different animal subjects, the distortion between frames becomes relatively severer.

We usually assume the stability prior during the calcium extraction step. Thus, an efficient motion-correction algorithm improves the accuracy of neuron extraction. Motion registration is often conducted by transforming each frame toward a reference frame using a mapping function. Linear transformations could be described by a linear matrix. Consider a 2D images as an example (3D volume could be generalized through adding an element to the coordinate vector). The mapping function between the coordinates  $x = [x_1, x_2, 1]^T$  and  $x' = [x'_1, x'_2, 1]^T$  of the corresponding sample points could be described as a  $3 \times 3$  linear transformation matrix  $T$ , where  $x' = Tx$ . The number of free parameters in  $T$  determines the transformation type, which could be divided into rigid transformation (translation and rotation) and non-rigid transformation (similarity, affine, and scaling)<sup>55</sup> [Fig. 2(a)]. Nonrigid or nonlinear transformations could also be realized by first splitting the image into overlapping patches to perform rigid correction, and merging the patches inversely for an partition-based non-rigid transformation field.<sup>57</sup> In head-fixed experiments, rigid transformations are usually sufficient for correction of a FOV smaller than 1 mm, and in mesoscale-imaging and freely moving animals, the non-rigid registrations are often demanded. 3D volume non-rigid registration raises the challenge of high computational load. The implementation of the graphic processing unit (GPU)-enhanced algorithms may be considered in large dataset registration missions. In some specific algorithms, the acceleration rate could be up to 100-fold.<sup>58</sup>

There are both inter-frame motion artifacts and intra-frame motion artifacts in neural imaging. The former exists widely in almost all kinds of microscopic modalities with the morphological shifting of the brain tissue. And the latter exists mainly in MPM because of the pixel-level temporal incoherence of sampling by point-scanning strategy [Fig. 2(b)]. Inter-frame motion artifacts are relatively more rigid than intra-frame motion artifacts, because there is usually no distortion within a single image, and a reference image is easier to find. A reference frame could be determined by visual inspection or averaging adjacent frames. All other frames could be registered to the reference frame.<sup>59</sup> The features used for registration could be pixel intensity,<sup>60</sup> extra structural channel,<sup>61</sup> or exogenous landmark.<sup>62</sup> A similarity measurements, such as pixel-wise difference, correlation, and information theory-based indexes, can also be used to evaluate



**Fig. 2** Inter-frame and intra-frame motion artifact correction. (a) Inter-frame motions are global transformations of the whole image which include rigid transformation (translation, rotation, and uniform scaling) and non-rigid transformation (scaling, affine, and projective). (b) Intra-frame motions render as pixel-wise displacement, which is usually caused by the motion during point scanning in multi-photon microscopies. (c) Diagram of intensity-based registration methodologies for inter-frame motion registration. (d) Example estimated trace produced by Lucas and Kanade algorithm, showing the displacement of the specimen in  $x$  (black) and  $y$  (cyan) directions in each frame and displacement difference between the two frames. Panels (b) and (d) are adapted from Ref. 56. Panel (c) is adapted from Ref. 55.

the difference between the reference frame and the moving frame. An extra regulation term of deformation was often included, to ensure the sparsity of the deformation field<sup>63</sup> and avoid over-fitting. Finally, various optimizers, mostly based on gradient descend theory, were designed to search for the best transformation parameter of the mapping function [Fig. 2(c)]. A thorough review of inter-frame motion artifact correction was given by Oliveira and Tavares.<sup>55</sup> Intra-frame motion artifacts were commonly corrected based on the Lucas and Kanade<sup>64</sup> algorithm [Fig. 2(d)] or a hidden Markov model,<sup>65,66</sup> where each pixel (or each line) was considered to be sampled from a translated tissue at an independent time-point.<sup>56,67</sup> The Lucas and Kanade algorithm expresses the difference between the registered frame and the template frame as the function of the  $x$ - $y$  trajectory. The optimal estimation of trajectory was derived by setting the first-order Taylor expansion of the error function to zero. And the algorithm iteratively updates the displacement in the  $x$ - $y$  axis, until it converges to an optimal  $x$ - $y$  displacement trace which minimizes the difference to the template. To reduce the computational cost, hierarchical approaches of image registration were proposed, where the imaging video was first decomposed into stable and non-stable sections, and different levels of registrations were assigned to the sectioned images.<sup>68</sup>

Apart from lateral motion, axial drifting eventually drew the researcher's attention with the evolution of increasing imaging time and resolution demand. In 2D parallel imaging,  $z$  axial motion correction could be accomplished by a hardware-implemented strategy of real-time focal plane adjustment,<sup>69</sup> or by computational approaches of calculating the correlation between the time-sequence  $z$  slices and a reference cube.<sup>61,65</sup> In high-speed volumetric imaging, the defocus problem is naturally resolved in a certain axial range because it could capture multi-focus planes at high speed. The  $z$  axis displacement could be equivalently resolved like the  $x$ - $y$  displacement by simply adding one more dimension to the registration pipeline. Since the axial displacement could not be avoided completely, volumetric imaging is much more robust to 3D motion artifacts compared with plane-scanning or point-scanning approaches.

A second image processing issue is denoising. Higher-SNR images are always preferred since it enhances the efficiency and fidelity of neuron detection and signal extraction. The simplest solution to acquire a higher-quality image was to use higher laser intensity. But it comes with photobleaching, nonlinear phototoxicity, and heating damage, which interferes with the fundamental neurophysiological phenomenon.<sup>70</sup> Meanwhile, it is worth noticing that under lower light conditions, photon shot noise becomes comparable to the readout noise of the camera sensor, where the hardware implementation could no longer eradicate noises. Therefore, to facilitate subsequent analysis, data-driven methods become critical.

Image denoising is routinely done by balancing the data fidelity and prior knowledge.<sup>71</sup> Commonly used priors in fluorescence imaging include sparsity prior,<sup>72</sup> also known as low-rank prior,<sup>73,74</sup> and sensor physics-based noise distribution prior.<sup>75</sup> Sparsity prior implies that spatial-temporal adjacent blocks in the image would have similar distributions. In other words, high-frequency components are dominated by unwanted noises and should be suppressed selectively. Different methods were used to exploit the sparsity priors. The most general method is to add a regularization term to the image reconstruction target function, which constrains the coefficient density in the transformation domain.<sup>76</sup> The deconvolution or image reconstruction optimization problem subsequently becomes a multi-target problem. The multiple targets could be decoupled under an alternating direction method of multipliers framework, leading to hybrid iterations of image reconstruction and denoising,<sup>73,77</sup> where the data fidelity term and the sparsity term are optimized iteratively. Within the sparsity sub-step, the parameters of the data fidelity term were fixed and the sparsity prior term was optimized, vice versa in the fidelity sub-step. There are various strategies to meet the sparsity prior. A straightforward approach is setting a threshold (soft or hard) cut-off in transformed domain<sup>72,75,78,79</sup> (Fourier transformation domain, cosine transformation domain, wavelet transformation domain, or a learned over-complete dictionary). The cut-off suppresses the high-frequency components and makes sure that the image frequency spectrum dominantly concentrates on low-frequency bands. Sparsity could also be attained using block-matching approaches. In block-matching-based algorithms, spatial-temporal adjacent blocks within the searching window were vectorized and concatenated into a matrix. The singular value decomposition was performed on the matrix, and then a hard or soft threshold was set

to suppress the low-energy components.<sup>80</sup> It was proved that this kind of strategy could efficiently reduce noise and preserve the details.<sup>72,78</sup>

Deep-learning methods have facilitated great advances in bioimage denoising.<sup>81</sup> Artificial neural networks were used to explore the underlying features and recover the noise-degraded images efficiently. There are two kinds of training strategies to train a network, supervised and unsupervised. The supervised strategy requires corresponding noisy and noise-free images to work as training data, and iteratively optimizes the network parameter using gradient descent strategy. In structural imaging *in vivo*, the training dataset could be acquired by pre-tests, where laser-insensitive specimens are used as subject.<sup>82</sup> And the pre-trained network could then be used to denoise the laser-sensitive specimen under low light conditions. While in functional imaging, this strategy might be problematic due to the non-repetitiveness of calcium transients, leading to the lack of training data for supervised learning. Therefore, self-supervised training was applied in functional neural image denoising issues. The network used solely noisy data as input and output in training. By utilizing the Noise2Noise framework<sup>83</sup> and temporal redundancy prior,<sup>84</sup> the network successfully produced noise-free images in test sessions. Thus, the self-supervised framework overcomes the obstacle of dataset shortage and is also practical in functional neural image missions.

Resolution maintenance is a vital concern in denoising algorithms designed for microscopy. Overall, the state of art frameworks such as local block-matching strategy and deep-learning network have less resolution loss compared with conventional transformational domain cut-off algorithms. But the performance is usually sample-specific and optical system-specific, and could not be easily summarized. The robustness is particularly a vital problem in deep-learning methods. Actually, there is a risk of resolution loss in every denoising algorithm because imaging denoising is essentially an ill-posed inverse problem. And there is a fundamental trade-off between detail preservation and denoising performance, in other words, the data fidelity and the prior knowledge. The ability of resolution preservation depends mainly on the accuracy of modeling the signal distribution and noise distribution, which requires the expertise on imaging system and sample property from researchers.

The following step of image preprocess is hemodynamic correction, which is typical in wide-field microscopy but is insignificant in two-photon microscopy. The blood flow in vessels has an impact on the background fluctuation of the neural image, which is exhibited as a large-variance and low-frequency background signal component. This component is calcium-irrelevant and should be removed. The most common approach to correct the hemodynamic is adding an extra reference channel with a specific excitation wavelength whose emission is calcium-independent.<sup>16,85</sup> The captured fluorescence intensity was divided by the reference channel,  $F_c = \frac{F}{R}$ , where  $F_c$  is the corrected signal;  $F$  is the calcium channel signal and  $R$  is the reference channel signal. The corrected normalized signal could be further expressed

as  $\Delta F_c = \frac{F_c - F_{c0}}{F_{c0}} = \frac{F/R}{F_0/R_0} - 1 = \frac{\frac{F_0(1+\Delta F)}{R_0}}{\frac{F_0(1+\Delta R)}{R_0}} - 1 = \frac{1+\Delta F}{1+\Delta R} - 1 \approx \Delta F - \Delta R$ , where subscript 0 stands

for averaged signal;  $\Delta F = \frac{F - F_0}{F_0}$ ,  $\Delta R = \frac{R - R_0}{R_0}$ ; the last equation is derived through a first order Taylor expansion. Another approach without extra reference channel is modeling the hemodynamic as background signal while performing calcium extraction. This computational approach relies on priors of calcium trace patterns and has no extra hardware cost. Detailed descriptions can be found in the next section.

## 2.2 Neuron Segmentation and Calcium Trace Extraction

After the image processing pipeline, the fluorescence images are low-noise and spatially settled and are ready for neuron segmentation and signal extraction. This session describes the methods used to extract fluorescence traces (or equivalent spike trains) and spatial footprints from the raw optical images. The data size should thus be reduced to several MBs which filters out the background fluorescence from neuropils, gliocytes, hemodynamic, and out-of-focus background while keeping all the neural-coding information intact. Compared with the raw optical video which might be up to hundreds of GBs or even TBs, this process could be regarded as data

compression or dimensionality reduction. Further neural functional and network inference could entirely rely on the extracted spatial-temporal footprint matrix.

An intuitive way of signal extraction is first segmenting individual neurons as regions of interest (ROI), and then calculating the (weighted) average temporal brightness of these pixels as the temporal trace. This kind of approach relies heavily on the imaging quality and meticulous identification of the shape and size of the neurons, which could be diverse between different bio-sensors and imaging modalities and often requires the expert knowledge of the researcher. Analytical approaches would assume neuron somas to be roughly circular in shape and flicker at a certain frequency periodically. Based on these assumptions, several image segmentation methods are used. First, the pixel-wise maximum deviation from the average brightness<sup>11,86</sup> or the standard deviation (SD) over time is derived to form an active map, from which neuron ROIs could be highlighted. Segmentation methods are subsequently used, ranging from manual approaches<sup>11,65</sup> to automatic algorithms. Manual ROI selection may be laborious but assures high ROI quality, which is suitable for small datasets. There are several software implementations available for manual ROI selections with the aid of automatic initializations, such as ImageJ,<sup>87</sup> SIMA,<sup>66</sup> and SamuROI.<sup>88</sup> Automatic algorithms for the segmentation are mainly based on computer vision theories, such as kernel filtering<sup>89</sup> and graph-cutting theory.<sup>66</sup> Deep-learning-based methods have also achieved state-of-art cell-segmentation performance,<sup>90-93</sup> while the conventional shortcomings of lack of training dataset, computational cost, and algorithm robustness to different imaging modalities and SNR levels are gradually increasing. Separation of overlapped neurons is a major challenge currently faced by deep-learning approaches. The anisotropic resolution further aggravated this problem. Different deep-learning methods adapted diverse strategies to tackle this issue. For instance, in U-Net<sup>91</sup> framework, the boundary between cells was artificially inserted into the mask of the training dataset, and the corresponding weight of the ridge was increased to force the network to learn the boundary, so overlapped cells were forced to be split into two non-overlap parts. In STNeuroNet,<sup>92</sup> overlapped neurons were split using watershed algorithm, and the temporal trace was demixed using a linear regression approach.<sup>94</sup> This would allow overlapping neurons to be separated spatially and temporally. And in Shallow U-Net Neuron Segmentation (SUNs),<sup>95</sup> segmentation was done frame-by-frame and followed by a merging procedure, in which case overlapped neurons firing at distinct frames could be separated. In the meantime, recent methods with multi-view imaging and reconstruction<sup>96</sup> can alleviate the anisotropic resolution, which may also increase the fidelity of segmentations.

The main drawback of the active-map-based segmentation is its high missing rate for inactive neurons and neurons with lower fluorescence indicator expression [Fig. 3(a)]. Another commonly used neural detection approach is matrix factorization. The core framework of these methods is to factorize the spatial-temporal matrix  $F \in \mathbb{R}^{N \times T}$  into spatial component  $S \in \mathbb{R}^{N \times K}$  and temporal component  $A \in \mathbb{R}^{K \times T}$  [Fig. 3(b)]. Each column of  $F$  containing  $N$  pixels is one of the  $T$  vectorized frames of the microscopic video. Ideally,  $S$  should contain  $K$  neurons' shapes and locations in every column. And  $A$  contains the corresponding temporal traces of these neurons. Additionally, there are also noise  $E$  and background signal  $B$  in  $F$ , which could be written as

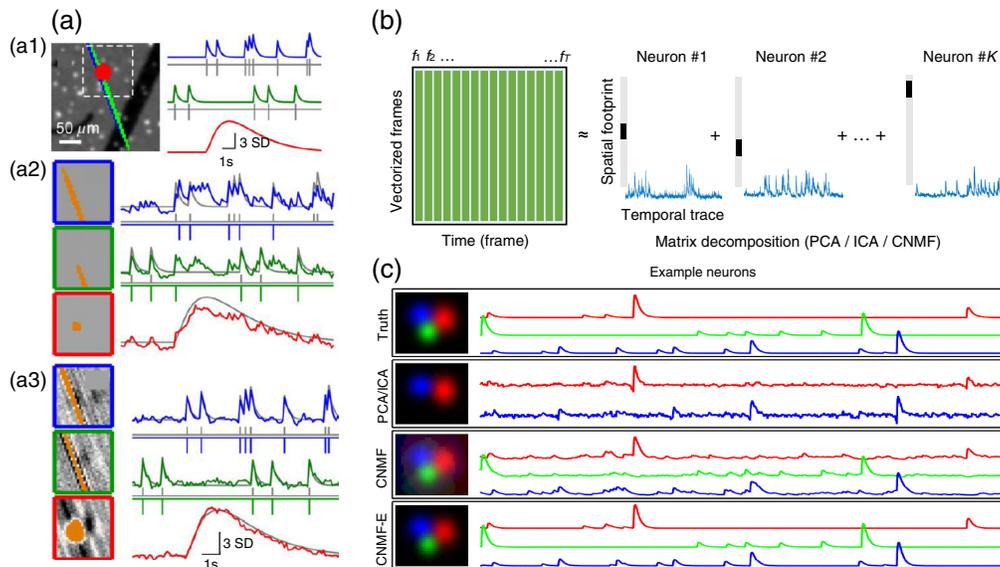
$$F = SA + B + E.$$

If assuming the spatial location of each neuron does not shift through time, which is facilitated by the previous motion registration process,  $SA$  could be further factorized into sum of  $K$  neuron's temporal traces. Thus  $SA$  could be represented as the sum of  $K$  rank-1 matrixes.

$$SA = \sum_{k=1}^K s_k a_k^T,$$

where  $s_k$  is the  $k$ 'th column of  $S$ , and  $a_k$  is the  $k$ 'th row of  $A$ .

There are infinite solutions to the factorization problem if non-constraints are set. If using only the least square of error as the criterion, the formula leads us to rank- $k$  approximation problem,



**Fig. 3** Calcium extraction via ROI-based approaches and matrix decomposition. (a) Comparison of ROI analysis and decomposition methods. The ground-truth responses are marked by gray lines. (a1) A simulated image which contains three spatially overlapping signal sources colored in red (Bergmann glial fibers), blue and green (Purkinje cell dendritic trees). The background contains black vessels and bright interneuron somata. (a2) ROI analysis identified spatial filters (left) of cellular components and their temporal traces (right). (a3) Matrix decomposition method reveals the spatial footprint (left) and their temporal traces (right). The signal estimation is more accurate than the ROI based-approach shown in (a2). (b) Schematic illustration of matrix decomposition methods on neural video. The video matrix was factorized into  $K$  rank-one matrixes, each of which stands for the location and temporal activity of an individual neuron. (c) The performance comparisons of PCA/ICA, CNMF, and CNMF-E on simulated data. Panel (a) is adapted from Ref. 97. Panel (c) is adapted from Ref. 98.

which is well solved by singular value decomposition, also known as principal component analysis (PCA). However, the PCA method alone might be ill-suited to extract single-neuron signals, since every principal component might contain information from multiple cells.<sup>97</sup> A further assumption is made about the statistical independence of each component, which leads to independent component analysis (ICA). ICA was proved to outperform PCA in identifying single-neuron cells.<sup>97</sup> Sparsity prior is another prior apart from independence. Sparse heterarchical matrix factorization<sup>99</sup> with dictionary learning is used to segment neurons and then cluster them into hierarchical functional clusters and reveal the network structure of functional circuits. Moreover, stricter priors of the shape and calcium response pattern could be made<sup>100</sup> which narrows down the solution space furthermore. Non-negative matrix factorization (NMF) method puts non-negative constraints on spatial and temporal matrix,<sup>30,101</sup> which is intuitive considering the non-negative nature of optical images and calcium activity traces. The NMF method performs better on noisy data compared with ICA and was widely adopted by a series of improved algorithms, such as constrained nonnegative matrix factorization (CNMF),<sup>102</sup> CNMF for microendoscope data (CNMF-E),<sup>98</sup> and CNMF with M-estimator to background<sup>103</sup> [Fig. 3(c)]. However, CNMF with 3D long-term mesoscale video dataset faces the burden of large computational costs up to thousands of GPU-hours.<sup>104</sup> Several accelerated algorithms are reported, such as seeded iterative demixing (SID)<sup>104</sup> and online deconvolution of calcium image.<sup>105,106</sup> Deep-learning algorithms can also be exploited in the future to further reduce the computational costs with more data priors.<sup>31</sup>

The fluorescence trace normally provides sufficient information as spike trains, since the two modalities are convertible through convolution and deconvolution. However, when precise timing information is wanted, deconvolution may be conducted to improve the data quality.<sup>54</sup> Under CNMF and OASIS framework, the binary spike train is embedded under the autoregressive model of calcium impulse response and could be achieved through primal optimization. If the

segmentation was done using ROI strategies, where spikes were not directly accessible, greedy template fitting was commonly used for spike event detection.<sup>8,107,108</sup>

Another common operation on the fluorescence trace is signal normalization, which extracts the mean activity of each neuron and calculates the difference from average score as  $\Delta F/F$ . These normalizations could further exclude the influence of incoherent background fluorescence and facilitate the analysis of neural functions.

### 3 Functional Mapping and Network Inference

At this phase of analysis, the information from fluorescence video is compressed into a 2D matrix  $X \in \mathbb{R}^{N \times T}$  containing the temporal traces of each neuron and the spatial maps of their locations on the brain atlas. The  $i$ 'th row of  $X$  represents the temporal trace of the  $i$ th neuron, lasting  $T$  frames in total. The information encoded in  $X$  is, to a large extent, similar to electrophysiological signals. The only difference is that optical signals reflect calcium transient intensity for calcium indicators (or other transient intensity for specific indicators such as voltage and other neurotransmitters), while electrophysiology signals reflect spike trains. These two modalities could be equivalently converted via event detection (deconvolution) and kernel convolution algorithm. Hence, the analysis methods used for these two modalities are generally consistent. In the following sections, we will introduce generic methods on functional mapping and network inference which would be applicable in both electrophysiological and fluorescence signal analysis, unless otherwise noted.

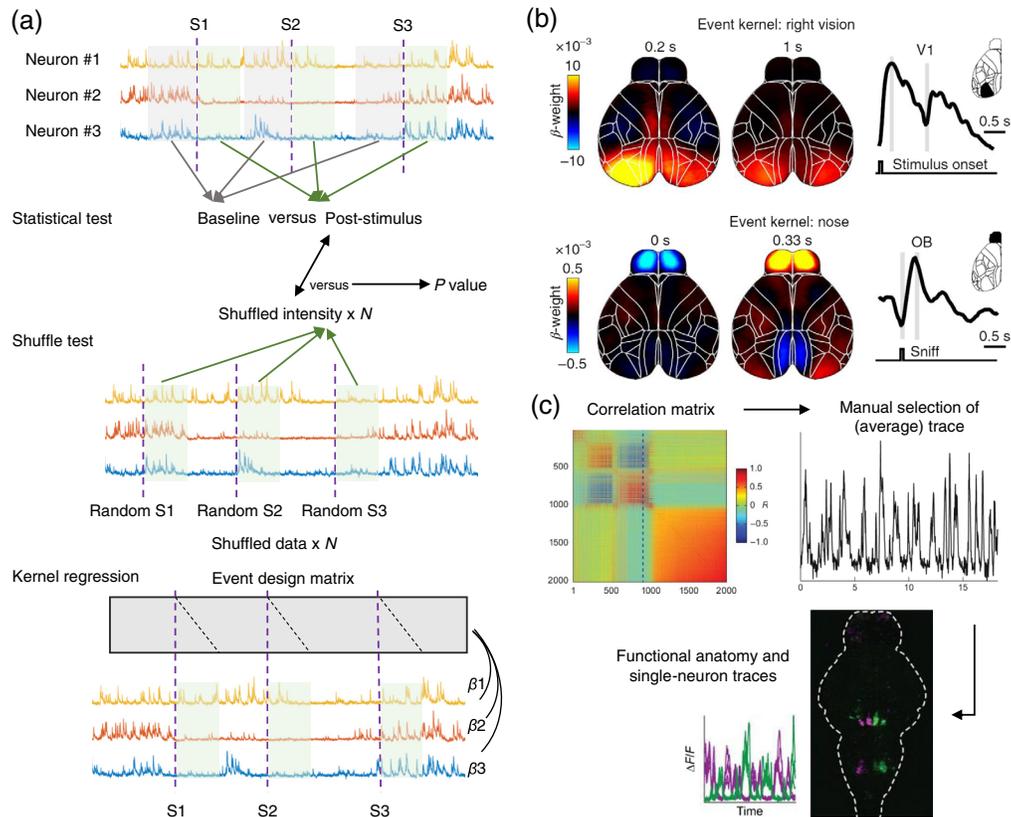
#### 3.1 Functional Mapping on Single-Neuron Resolution

An emerging challenge for mesoscale optical recording of neural signals is that task-relevant neurons only take up a small proportion of captured neurons and might be spatially dispersed over the FOV.<sup>109</sup> Most of the neurons over the cortex encode spontaneous and un instructed movements unrelated to the task.<sup>110</sup> Including all these neurons in analyses would cost large computational costs and introduce undesirable noises, which reduce the SNR of the neural encoding space. Thus, a first and foremost step before functional analysis is to discriminate the task-relevant neurons from the massive number of entire neurons.

A forthright intuition is to select neurons that exhibited a more active firing pattern during the task [Fig. 4(a)]. The degree to which task events modulate neuron activities could be evaluated through statistical significance tests by comparing the averaged firing rate between task trials and baseline.<sup>109</sup> Diverse time windows could be used, such as pre-stimulus, post-stimulus, and post-movement onset, depending on the interested task variable. However, the significance test employs nothing on the unique distribution of individual neuron activity prior. Neurons with periodic firing or with extremely active or inactive firing properties may be misjudged by statistical tests. Thus, an alternative option is to perform a shuffle test on the same neuron's activity trace. Shuffling the start time stamp of behavior segments<sup>11</sup> or shuffling the timepoint of calcium events<sup>8</sup> are both feasible. The same firing rate average analysis is performed on the shuffled data. The significance of task modulation is evaluated through how many percentiles of the shuffled data the primary data could exceed.

Significance tests revealed the overall task modulation on the firing rate during a task period. But within a complicated task, there might be more than one task variable, such as engagement, choice, precision, and body gestures. Statistical tasks based on mean firing rate perform badly on separating these fine variables or revealing the non-linear neural representation. To uncover the relationship between single-neuron activity and animal behaviors, superior data analysis methods should be employed.

Linear regression, also known as general linear model is an effective method for exploring data correlations. The regression model could be expressed as  $y = X\beta$ ,<sup>108</sup> where  $\beta$  is a concatenated event kernel vector to be regressed.  $X \in \mathbb{R}^{T \times L}$  is a Toeplitz design matrix concatenated by  $K$  events of interest  $X_k$ , where  $X_k \in \mathbb{R}^{T \times L_k}$ ,  $\sum_{k=1}^K L_k = L$ ,  $X = [X_1, X_2, \dots, X_K]$ .  $X_k$  contains diagonal 1s at each event onset and corresponding time lags and 0 elsewhere.  $y$  is the normalized neuron trace time course as the prediction output. The link strength between the neuron trace and



**Fig. 4** Examples of functional mapping and functional circuit identification. (a) Schematic of task-relevant neuron evaluation. About three example neuron traces are shown with three stimuli as event triggers (dashed lines). In statistical test approaches (top), the fluorescence intensity in the pre-stimulus baseline window (shaded in gray) is compared with post-stimulus time windows (shaded in green), deriving the significance index. In shuffle test approaches (middle), post-stimulus reactions in the shuffled dataset are compared with observed true reaction, where a  $P$ -value is defined as the percentile of exceeded shuffled data by the observed data. In the kernel regression (bottom) method, the linear model was used to predict the neuron response from the event design matrix. The link between the task and the neuron activity is evaluated via the regression coefficient  $\beta$ . (b) Example weight maps of event kernel for right visual stimulus and nose movement at pixel-level. (c) Diagram of correlation-based functional clustering at the voxel level. The correlation matrix of the pre-selected supervoxels is manually inspected and a reference trace is selected. Each voxel in the brain is correlated to the reference trace, forming a correlation map. Green, positive correlations; magenta, negative correlations. Panel (b) is adapted from Ref. 110. Panel (c) is adapted from Ref. 60.

the behavior variable could be evaluated through the coefficient in  $\beta$ . A larger absolute value in  $\beta$  stands for a higher correlation between neuron and behavior. The spatial-temporal map of  $\beta$  could also indicate the information flow between different neuron clusters. To avoid overfitting, different regulation terms are added to rescript the sparsity of the coefficient, which make the model more practical, such as reduced-rank kernel regression,<sup>109</sup> ridge regression<sup>108</sup> or lasso regression<sup>110</sup> [Fig. 4(b)].

In perceptual pathways, such as tactile, olfaction, and vision, neural coding mechanisms are often better explored, thus allowing more accurate and complicated modeling of neural activities. For instance, in the tactile sensory pathway, the functional mapping modal assumes that the sensory input goes past static, point-wise nonlinearity, and is convolved with a temporal kernel before being input to a Gaussian noise model.<sup>8</sup> While in the olfaction pathway, different olfactory is believed to be encoded by population vector instead of single-neuron activities.<sup>9</sup> And in the visual pathway, the ganglion neurons are well known to be modulated by surrounding neurons, forming the widely recognized receptive field.<sup>111</sup> These models successfully predict the activity in the local circuit with high accuracy and specificity.

Comparing the modal-free analysis with the model-based analysis, the former is more helpful for rough neuron sorting from mesoscale images, while the latter uncovers the explainable mechanism of neural encoding. Thus, we advocate taking advantage of the high data-throughput property of mesoscale imaging as a neural information database, from which neuron clusters could be distinguished and separated depending on their functions. While the subsequent modeling of neural encoding could be more targeted and prior-based. This may extend the boundary of controversial single-brain-area functional mapping restriction.

### 3.2 Brain-Wide Single-Neural Level Functional Network Study

It is widely recognized that different levels of clusters of neurons are recruited for diverse behavior tasks.<sup>112</sup> Even a simple perceptual input could arouse distributed neural activities across the brain. Communication across multiple regions is crucial for the brain functioning as a system.<sup>29,109,113</sup> Observing the brain *in vivo* with behaving animals enables the inference of functional circuits across brain areas, by analyzing statistical dependencies between neurons.

Restricted by the detection capability, conventional network inference either concentrated on a local circuit or utilized low-spatial-resolution brain activities through functional magnetic resonance imaging,<sup>113,114</sup> electroencephalogram, or wide-field microscopy.<sup>29</sup> Brain activities acquired through these techniques characterize the area-averaged multi-unit neuron activity containing multi-frequency oscillation components,<sup>115</sup> thus are temporally smoother and more stable across trials. Single-neuron activities, however, exhibited larger variance, sparsity, and randomness through trials.<sup>116</sup> The commonly used spectrum methods on field potentials decompose different frequency components and analyze the phase lag between brain areas to infer the information flow through the precedence of brain areas. These methods became meaningless on single-neuron traces because the frequency components lack interpretable physical meanings. Metrics of similarity such as instantaneous phase, phase lag, and phase synchronization index<sup>117</sup> are no longer applicable. And the quadratic increase of the computational complexity while calculating pair-wise similarity of the neurons also brings challenges to the network inference. All these barriers bring great difficulties and opportunities to interpret the working mechanisms of neural circuits at the single-neuron level with mesoscale neural imaging.

Despite the inefficiency of phase decomposition-based methods, the covariance or correlation remains the most straightforward model-based metric which describes the undirected similarity between two fluorescence signals. The similarity matrix represents the pairwise correlation among the neural population. With the pairwise correlation which describes the similarity of neurons, subsequent clustering<sup>118</sup> methods could be used to assemble the neurons into separated groups.<sup>119</sup> Manually selected seed with Pearson correlation matrix has been used to identify the functional circuits across the larval zebrafish brain<sup>60</sup> [Fig. 4(c)]. There are plentiful other unsupervised clustering methods available, such as *k*-means clustering, hierarchical clustering, barcode analysis, and graph-based analysis. These methods have been used in different experiments. The *k*-means clustering was used to find the representation structure at single-neuron level in the rat orbitofrontal cortex.<sup>120</sup> The *k*-means method remains computationally efficient even with high dimensional data, but the result of *k*-means algorithms depends heavily on the hyperparameter *k*, thus requiring pretest or the prior knowledge from the researcher. The hierarchical algorithm was used to cluster the retinal ganglion cells and exhibited high interpretability and visualization ability.<sup>121</sup> However, it may encounter severe computational complexity with higher-dimensional data. The barcode analysis required a character barcoding process before categorizing. For instance, the neurons from the whole brain of the zebrafish were barcoded via their response to certain stimuli and clustered into 256 classes.<sup>86</sup> What may be troublesome with barcoding is that the number of categories increases exponentially with the bit of barcode in binary coding, which restricts its application scope.

Apart from the model-based clustering, functional cell assemblies could also be identified through model-free decomposition analysis. Factorization methods are performed on the 3D tensor stacked by multi-trial 2D neural activity maps, which includes PCA, demixed principal component analysis (demixed PCA),<sup>122</sup> and non-negative tensor factorization.<sup>123</sup> These models factorize the tensor into rank-1 components and extract the temporal and trial dimensional vector as trial-consistent and within-trial latent variables of the neuron representation. Though

unsupervised, these vectors may correspond to behavioral observations such as performance accuracy and task engagement.

The temporal correlation-based network mentioned above often induces false-positive connections between indirectly coupled neurons.<sup>112</sup> A sparser binary adjacency matrix could be derived by hard-threshold or  $k$ -nearest neighbors. On the other hand, Granger causality, or transfer entropy analysis, is a temporal-precedency-based metric that reveals directed information flow from one node to another. Its basic assumption is that the history of a precedent neuron should be contributing to predicting the future of the downstream neuron. Granger causality has been used to investigate the unbalanced distribution of information density in the somatosensory cortex of the mouse.<sup>124</sup> Additionally, there has been open-source algorithm toolboxes<sup>125</sup> for performing Granger causality analysis. The computational cost of Granger causality algorithm increases quadratically with the number of nodes. Improved Granger causality algorithms<sup>126</sup> are developed to decrease computational costs and avoid overfitting. Transfer entropy also serves as an equivalent for Granger causality under Gaussian variables<sup>127</sup> with a much lower computational complexity. With the help of optogenetics,<sup>128</sup> it is possible to perturbate *in vivo* neurons and watch the network in dynamic. Modal-based Bayesian network was used to inference the neural spike trains and connectivity with simultaneous optogenetic perturb and electrophysiological recording.<sup>129</sup>

## 4 Looking Ahead

Mesoscale imaging with single-neuron level resolution has provided unprecedented potential for understanding the mechanism of both local and long-range brain circuits. Ambitious computational scientists look forward to mimicking the brain functions and building general intelligent machines. But there are still data analysis obstacles lying ahead of us, which would take long-term efforts before we uncover the myths of the brain network.

The fast-growing of long-term high-speed mesoscale volumetric imaging craves high-efficiency data analysis methods. The volume sequences captured by light-sheet microscopes, light-field microscopes, or multifocal microscopes could easily make TB-level data sizes. For a typical example, conventional matrix decomposition approaches are ineffectual for calcium extraction from 3D volumes. SID<sup>104</sup> was used to reconstruct 3D positions of neurons by first deconvolving the pixel-wise SD image of background-subtracted images. The neuron candidate positions were identified using a band-pass filter and back-projected to multiple views for spatial-temporal footprint update. This has enabled the decomposition to take place in a lower dimension, which saves vastly the computational cost. Further analysis considering the pairwise correlation or causality between neurons also costs massive computing resources. Apart from the task-relevant sorting procedure, which could massively reduce the analyzing subjects, dimensionality reduction under sparsity prior<sup>118</sup> could also help us with identifying hierarchical clusters of neuron assemblies. Nevertheless, analyzing the populational property of neurons increases the stability over trial and time-lapse, and surrenders the single-neuron level calculation mechanism, which should be a trade-off that requires careful consideration. Data-driven deep-learning methods are very promising in many steps of the whole pipeline especially for its low computational costs, since the mesoscale imaging data usually has very strong local similarity across a large FOV. The whole pipeline should be considered simultaneously during the design of the network framework, while the downstream analysis can also be used as the metric to evaluate the performance of different algorithms. The generalization and practicality of current deep neural networks need to be enhanced for its broad and convenient applications. In addition, more and more mesoscale databases of different imaging modalities and diverse tasks are required to promote this field and evaluate the rapidly emerging algorithms.

Another concern of single-neuron network study, as well as higher-level networks inference of brains, is the network interpretability.<sup>130</sup> There are researchers arguing the credibility of the commonly used methods such as single-neuron lesion, correlation, and granger causality.<sup>131</sup> The analytic approaches were performed on a microprocessor whose algorithm flowchart was known as prior. The research concluded pessimistically that current approaches may fall short of interpreting the algorithm of the neural system, regardless of the amount of data. It seems that although we are desperate for acquiring more data, the basic tool of understanding the brain

network is still absent. It was suggested by Marr<sup>132</sup> that there are three levels of understanding a system: functions, algorithms, and implementation. Although we have dissected several perceptual input pathways such as vision and audition, the basic communication protocol of neurons remained unknown. To understand how the neurons in brain give rise to the ability of recognition, we may need a network analysis technique that could find more fundamental arithmetic units or their hierarchical structures of neural network.

In conclusion, we have introduced the general analysis pipeline of fluorescence mesoscale brain images in this review. But there are still multiple challenges for existing methods to deal with the complex network of the brain. Nevertheless, we believe that with the rapid growth of imaging techniques, analyzing methods, and computational abilities, global circuits on the single-neuron level will be more extensively explored in the future.

## Disclosures

The authors declare no conflict of interests.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 62088102 and 62071272) and the National Key Research and Development Program of China (Grant No. 2020AA0105500). We further thank the supports from Beijing Laboratory of Brain and Cognitive Intelligence, Beijing Municipal Education Commission, and Beijing Key Laboratory of Multi-Dimension & Multi-Scale Computational Photography (MMCP).

## Code, Data and Materials Availability

Example data and code in this paper are currently not publicly available, but may be obtained from the author upon reasonable request.

## References

1. M. Scanziani and M. Hausser, "Electrophysiology in the age of light," *Nature* **461**(7266), 930–939 (2009).
2. T. H. Kim and M. J. Schnitzer, "Fluorescence imaging of large-scale neural ensemble dynamics," *Cell* **185**(1), 9–41 (2022).
3. T. H. Kim et al., "Long-term optical access to an estimated one million neurons in the live mouse cortex," *Cell Rep.* **17**(12), 3385–3394 (2016).
4. I. V. Kauvar et al., "Cortical observation by synchronous multifocal optical sampling reveals widespread population encoding of actions," *Neuron* **107**(2), 351 (2020).
5. J. Cichon and W.-B. Gan, "Branch-specific dendritic Ca<sup>2+</sup> spikes cause persistent synaptic plasticity," *Nature* **520**(7546), 180–185 (2015).
6. C. S. W. Lai, T. F. Franke, and W.-B. Gan, "Opposite effects of fear conditioning and extinction on dendritic spine remodelling," *Nature* **483**(7387), 87–91 (2012).
7. M. H. Mohajerani et al., "Spontaneous cortical activity alternates between motifs defined by regional axonal projections," *Nat. Neurosci.* **16**, 1426–1435 (2013).
8. S. P. Peron et al., "A cellular resolution map of barrel cortex activity during tactile behavior," *Neuron* **86**(3), 783–799 (2015).
9. C. E. Schoonover et al., "Representational drift in primary olfactory cortex," *Nature* **594**(7864), 541–546 (2021).
10. T.-W. Chen et al., "A map of anticipatory activity in mouse motor cortex," *Neuron* **94**(4), 866 (2017).
11. A. Klaus et al., "The spatiotemporal organization of the striatum encodes action space," *Neuron* **96**(4), 949 (2017).
12. L. N. Driscoll et al., "Dynamic reorganization of neuronal activity patterns in parietal cortex," *Cell* **170**(5), 986 (2017).

13. J. G. Parker et al., “Diametric neural ensemble dynamics in parkinsonian and dyskinetic states,” *Nature* **557**(7704), 177–182 (2018).
14. D. Mao et al., “Sparse orthogonal population representation of spatial context in the retrosplenial cortex,” *Nat. Commun.* **8**, 243 (2017).
15. J. A. Berry, A. Phan, and R. L. Davis, “Dopamine neurons mediate learning and forgetting through bidirectional modulation of a memory trace,” *Cell Rep.* **25**(3), 651–662.e5 (2018).
16. W. E. Allen et al., “Global representations of goal-directed behavior in distinct cell types of mouse neocortex,” *Neuron* **94**(4), 891 (2017).
17. J. Fan et al., “Video-rate imaging of biological dynamics at centimetre scale and micrometre resolution,” *Nat. Photon.* **13**, 809–816 (2019).
18. R. Lu et al., “Rapid mesoscale volumetric imaging of neural activity with synaptic resolution,” *Nat. Methods* **17**, 291–294 (2020).
19. J. Demas et al., “High-speed, cortex-wide volumetric recording of neuroactivity at cellular resolution using light beads microscopy,” *Nat. Methods* **18**, 1103–1111 (2021).
20. J. Wu et al., “Iterative tomography with digital adaptive optics permits hour-long intravital observation of 3D subcellular dynamics at millisecond scale,” *Cell* **184**, 3318–3332.e17 (2021).
21. Z. Zhang et al., “Imaging volumetric dynamics at high speed in mouse and zebrafish brain with confocal light field microscopy,” *Nat. Biotechnol.* **39**, 74–83 (2021).
22. F. F. Voigt et al., “The mesoSPIM initiative: open-source light-sheet microscopes for imaging cleared tissue,” *Nat. Methods* **16**, 1105–1108 (2019).
23. A. San Martin et al., “Single-cell imaging tools for brain energy metabolism: a review,” *Neurophotonics* **1**(1), 011004 (2014).
24. W. Yang and R. Yuste, “In vivo imaging of neural activity,” *Nat. Methods* **14**, 349–359 (2017).
25. S. Weisenburger and A. Vaziri, “A guide to emerging technologies for large-scale and whole-brain optical imaging of neuronal activity,” *Annu. Rev. Neurosci.* **41**, 431–452 (2018).
26. M. Z. Lin and M. J. Schnitzer, “Genetically encoded indicators of neuronal activity,” *Nat. Neurosci.* **19**, 1142–1153 (2016).
27. R. Y. Tsien, “New calcium indicators and buffers with high selectivity against magnesium and protons: design, synthesis, and properties of prototype structures,” *Biochemistry* **19**(11), 2396–2404 (1980).
28. Y. Wu et al., “Advanced optical imaging techniques for neurodevelopment,” *Curr. Opin. Neurobiol.* **23**(6), 1090–1097 (2013).
29. C. Ren and T. Komiyama, “Characterizing cortex-wide dynamics with wide-field calcium imaging,” *J. Neurosci.* **41**(19), 4160–4168 (2021).
30. C. Moretti and S. Gigan, “Readout of fluorescence functional signals through highly scattering tissue,” *Nat. Photon.* **14**, 361–364 (2020).
31. L. Sita et al., “A deep-learning approach for online cell identification and trace extraction in functional two-photon calcium imaging,” *Nat. Commun.* **13**, 1529 (2022).
32. O. A. Shemesh et al., “Precision calcium imaging of dense neural populations via a cell-body-targeted calcium indicator,” *Neuron* **107**(3), 470–486.e11 (2020).
33. Y. Chen et al., “Soma-targeted imaging of neural circuits by ribosome tethering,” *Neuron* **107**(3), 454–469.e6 (2020).
34. J. Jonkman et al., “Tutorial: guidance for quantitative confocal microscopy,” *Nat. Protoc.* **15**, 1585–1611 (2020).
35. R. M. Power and J. Huisken, “A guide to light-sheet fluorescence microscopy for multi-scale imaging,” *Nat. Methods* **14**, 360–373 (2017).
36. X. Y. Deng and M. Gu, “Penetration depth of single-, two-, and three-photon fluorescence microscopic imaging through human cortex structures: Monte Carlo simulation,” *Appl. Opt.* **42**(16), 3321–3329 (2003).
37. K. Svoboda and R. Yasuda, “Principles of two-photon excitation microscopy and its applications to neuroscience,” *Neuron* **50**(6), 823–839 (2006).
38. T. Wang and C. Xu, “Three-photon neuronal imaging in deep mouse brain,” *Optica* **7**(8), 947–960 (2020).

39. F. Helmchen and W. Denk, “Deep tissue two-photon microscopy,” *Nat. Methods* **2**, 932–940 (2005).
40. T. A. Murray and M. J. Levene, “Singlet gradient index lens for deep *in vivo* multiphoton microscopy,” *J. Biomed. Opt.* **17**(2), 021106 (2012).
41. N. Ji, J. Freeman, and S. L. Smith, “Technologies for imaging neural activity in large volumes,” *Nat. Neurosci.* **19**, 1154–1164 (2016).
42. V. Iyer, T. M. Hoogland, and P. Saggau, “Fast functional imaging of single neurons using random-access multiphoton (RAMP) microscopy,” *J. Neurophysiol.* **95**(1), 535–545 (2006).
43. W. Amir et al., “Simultaneous imaging of multiple focal planes using a two-photon scanning microscope,” *Opt. Lett.* **32**(12), 1731–1733 (2007).
44. D. R. Beaulieu et al., “Simultaneous multiplane imaging with reverberation two-photon microscopy,” *Nat. Methods* **17**, 283–286 (2020).
45. J. Wu et al., “Kilohertz two-photon fluorescence microscopy imaging of neural activity *in vivo*,” *Nat. Methods* **17**, 287–290 (2020).
46. Y. Xue et al., “Single-shot 3D wide-field fluorescence imaging with a computational miniature mesoscope,” *Sci. Adv.* **6**(43) (2020).
47. M. Levoy et al., “Light field microscopy,” *ACM Trans. Graphics* **25**(3), 924–934 (2006).
48. R. Prevedel et al., “Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy,” *Nat. Methods* **11**, 727–730 (2014).
49. M. Broxton et al., “Wave optics theory and 3-D deconvolution for the light field microscope,” *Opt. Express* **21**(21), 25418–25439 (2013).
50. Y. Zhang et al., “Computational optical sectioning with an incoherent multiscale scattering model for light-field microscopy,” *Nat. Commun.* **12**, 6391 (2021).
51. B. B. Scott et al., “Imaging cortical dynamics in GCaMP transgenic rats with a head-mounted widefield macroscope,” *Neuron* **100**(5), 1045–1058.e5 (2018).
52. J. Couto et al., “Chronic, cortex-wide imaging of specific cell populations during behavior,” *Nat. Protoc.* **16**, 3241–3263 (2021).
53. E. A. Pnevmatikakis, “Analysis pipelines for calcium imaging data,” *Curr. Opin. Neurobiol.* **55**, 15–21 (2019).
54. C. Stringer and M. Pachitariu, “Computational processing of neural recordings from calcium imaging data,” *Curr. Opin. Neurobiol.* **55**, 22–31 (2019).
55. F. P. M. Oliveira and J. M. R. S. Tavares, “Medical image registration: a review,” *Comput. Methods Biomech. Biomed. Eng.* **17**(2), 73–93 (2014).
56. D. S. Greenberg and J. N. D. Kerr, “Automated correction of fast motion artifacts for two-photon imaging of awake animals,” *J. Neurosci. Methods* **176**(1), 1–15 (2009).
57. E. A. Pnevmatikakis and A. Giovannucci, “NoRMCorre: an online algorithm for piecewise rigid motion correction of calcium imaging data,” *J. Neurosci. Methods* **291**, 83–94 (2017).
58. M. Guo et al., “Rapid image deconvolution and multiview fusion for optical microscopy,” *Nat. Biotechnol.* **38**, 1337–1346 (2020).
59. D. Soulet et al., “Automated filtering of intrinsic movement artifacts during two-photon intravital microscopy,” *PLOS ONE* **8**(1), e53942 (2013).
60. M. B. Ahrens et al., “Whole-brain functional imaging at cellular resolution using light-sheet microscopy,” *Nat. Methods* **10**, 413–420 (2013).
61. T. M. Ryan et al., “Correction of *z*-motion artefacts to allow population imaging of synaptic activity in behaving mice,” *J. Physiol.-Lond.* **598**(10), 1809–1827 (2020).
62. V. A. Griffiths et al., “Real-time 3D movement correction for two-photon imaging in behaving animals,” *Nat. Methods* **17**, 741–748 (2020).
63. G. Auzias et al., “Diffeomorphic brain registration under exhaustive sulcal constraints,” *IEEE Trans. Med. Imaging* **30**(6), 1214–1227 (2011).
64. B. D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” in *Proc. 7th Int. Joint Conf. Artif. Intell.*, Morgan Kaufmann Publishers Inc., Vancouver, British Columbia, Vol. 2, pp. 674–679 (1981).
65. D. A. Dombeck et al., “Imaging large-scale neural activity with cellular resolution in awake, mobile mice,” *Neuron* **56**(1), 43–57 (2007).

66. P. Kaifosh et al., “SIMA: Python software for analysis of dynamic fluorescence imaging data,” *Front. Neuroinf.* **8**, 80 (2014).
67. N. S. Alexander et al., “Image registration and averaging of low laser power two-photon fluorescence images of mouse retina,” *Biomed. Opt. Express* **7**(7), 2671–2691 (2016).
68. J. Lu et al., “MINIPIPE: a miniscope 1-photon-based calcium imaging signal extraction pipeline,” *Cell Rep.* **23**(12), 3673–3684 (2018).
69. J. L. Chen et al., “Online correction of licking-induced brain motion during two-photon imaging with a tunable lens,” *J. Physiol.* **591**(19), 4689–4698 (2013).
70. J. Icha et al., “Phototoxicity in live fluorescence microscopy, and how to avoid it,” *Bioessays* **39**(8), (2017).
71. S. Gu and R. Timofte, “A brief review of image denoising algorithms and beyond,” in *Inpainting and Denoising Challenges*, S. Escalera et al., Ed., pp. 1–21, Springer International Publishing, Cham (2019).
72. K. Dabov et al., “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
73. Y. Liu et al., “Rank minimization for snapshot compressive imaging,” *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(12), 2990–3006 (2019).
74. J. He et al., “Spatial-temporal low-rank prior for low-light volumetric fluorescence imaging,” *Opt. Express* **29**(25), 40721–40733 (2021).
75. B. Mandracchia et al., “Fast and accurate sCMOS noise correction for fluorescence microscopy,” *Nat. Commun.* **11**, 94 (2020).
76. N. Dey, “Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution,” *Microsc. Res. Tech.* **69**, 260–266 (2006).
77. W. Dong, “Compressive sensing via nonlocal low-rank regularization,” *IEEE Trans. Image Process.* **23**, 3618–3632 (2014).
78. M. Alain and A. Smolic, “Light field denoising by sparse 5D transform domain collaborative filtering,” in *IEEE 19th Int. Workshop Multimedia Signal Process.*, pp. 1–6 (2017).
79. Y. Zhang et al., “DiLFM: an artifact-suppressed and noise-robust light-field microscopy through dictionary learning,” *Light-Sci. Appl.* **10**(1), 152 (2021).
80. J.-F. Cai, E. J. Candes, and Z. Shen, “A singular value thresholding algorithm for matrix completion,” *Siam J. Optim.* **20**(4), 1956–1982 (2010).
81. R. F. Laine, G. Jacquemet, and A. Krull, “Imaging in focus: an introduction to denoising bioimages in the era of deep learning,” *Int. J. Biochem. Cell Biol.* **140**, 106077 (2021).
82. M. Weigert et al., “Content-aware image restoration: pushing the limits of fluorescence microscopy,” *Nat. Methods* **15**, 1090–1097 (2018).
83. J. Lehtinen et al., “Noise2Noise: learning image restoration without clean data,” in *Proc. Mach. Learn. Res., 35th Int. Conf. Mach. Learn.* (2018).
84. X. Li et al., “Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising,” *Nat. Methods* **18**, 1395–1400 (2021).
85. J. B. Wekselblatt et al., “Large-scale imaging of cortical dynamics during sensory perception and behavior,” *J. Neurophysiol.* **115**(6), 2852–2866 (2016).
86. E. A. Naumann et al., “From whole-brain data to functional circuit models: the zebrafish optomotor response,” *Cell* **167**(4), 947–960.e20 (2016).
87. C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, “NIH image to ImageJ: 25 years of image analysis,” *Nat. Methods* **9**, 671–675 (2012).
88. M. Rueckl et al., “SamuROI, a Python-based software tool for visualization and analysis of dynamic time series imaging at multiple spatial scales,” *Front. Neuroinf.* **11**, 1–14 (2017).
89. S. L. Smith and M. Hausser, “Parallel processing of visual space by neighboring neurons in mouse visual cortex,” *Nat. Neurosci.* **13**(9), 1144–1149 (2010).
90. C. Wachinger, M. Reuter, and T. Klein, “DeepNAT: deep convolutional neural network for segmenting neuroanatomy,” *Neuroimage* **170**, 434–445 (2018).
91. T. Falk et al., “U-Net: deep learning for cell counting, detection, and morphometry,” *Nat. Methods* **16**, 67–70 (2019).

92. S. Soltanian-Zadeh et al., “Fast and robust active neuron segmentation in two-photon calcium imaging using spatiotemporal deep learning,” *Proc. Natl. Acad. Sci. U. S. A.* **116**(17), 8554–8563 (2019).
93. Z. Gu et al., “CE-Net: context encoder network for 2D medical image segmentation,” *IEEE Trans. Med. Imaging* **38**(10), 2281–2292 (2019).
94. S. E. J. de Vries et al., “A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex,” *Nat. Neurosci.* **23**, 138–151 (2020).
95. Y. J. Bao et al., “Segmentation of neurons from fluorescence calcium recordings beyond real time,” *Nat. Mach. Intell.* **3**, 590–600 (2021).
96. Y. Wu et al., “Multiview confocal super-resolution microscopy,” *Nature* **600**(7888), 279–284 (2021).
97. E. A. Mukamel, A. Nimmerjahn, and M. J. Schnitzer, “Automated analysis of cellular signals from large-scale calcium imaging data,” *Neuron* **63**(6), 747–760 (2009).
98. P. Zhou et al., “Efficient and accurate extraction of *in vivo* calcium signals from microendoscopic video data,” *Elife* **7**, e28728 (2018).
99. F. Diego and F. A. Hamprecht, “Learning multi-level sparse representations,” in *Proc. 26th Int. Conf. Neural Inf. Process. Syst.*, Curran Associates Inc., Lake Tahoe, Nevada, Vol. 1, pp. 818–826 (2013).
100. F. Diego and F. A. Hamprecht, “Sparse space-time deconvolution for Calcium image analysis,” in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, MIT Press, Montreal, Vol. 1, pp. 64–72 (2014).
101. R. Maruyama et al., “Detecting cells using non-negative matrix factorization on calcium imaging data,” *Neural Networks* **55**, 11–19 (2014).
102. E. A. Pnevmatikakis et al., “Simultaneous denoising, deconvolution, and demixing of calcium imaging data,” *Neuron* **89**(2), 285–299 (2016).
103. H. Inan, M. A. Erdogdu, and M. J. Schnitzer, “Robust estimation of neural signals in calcium imaging,” in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, Curran Associates Inc., Long Beach, California, pp. 2905–2914 (2017).
104. T. Nobauer et al., “Video rate volumetric Ca<sup>2+</sup> imaging across cortex using seeded iterative demixing (SID) microscopy,” *Nat. Methods* **14**(8), 811–818 (2017).
105. J. Friedrich, P. Zhou, and L. Paninski, “Fast online deconvolution of calcium imaging data,” *PLoS Comput. Biol.* **13**(3), e1005423 (2017).
106. A. Goyannucci et al., “OnACID: online analysis of calcium imaging data in real time,” in *31st Annu. Conf. Neural Inf. Process. Syst.* (2017).
107. H. Lütke et al., “Inference of neuronal network spike dynamics and topology from calcium imaging data,” *Front. Neural Circuits* **7**, 201 (2013).
108. A. J. Peters et al., “Striatal activity topographically reflects cortical activity,” *Nature* **591**(7850), 420–425 (2021).
109. N. A. Steinmetz et al., “Distributed coding of choice, action and engagement across the mouse brain,” *Nature* **576**(7786), 266 (2019).
110. S. Musall et al., “Single-trial neural dynamics are dominated by richly varied movements,” *Nat. Neurosci.* **22**, 1677–1686 (2019).
111. J. W. Pillow et al., “Spatio-temporal correlations and visual signalling in a complete neuronal population,” *Nature* **454**(7207), 995–999 (2008).
112. D. S. Bassett and O. Sporns, “Network neuroscience,” *Nat. Neurosci.* **20**, 353–364 (2017).
113. F. V. Farahani, W. Karwowski, and N. R. Lighthall, “Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review,” *Front. Neurosci.* **13**, 585 (2019).
114. M. W. Cole et al., “Intrinsic and task-evoked network architectures of the human brain,” *Neuron* **83**(1), 238–251 (2014).
115. A. M. Bastos and J.-M. Schoffelen, “A tutorial review of functional connectivity analysis methods and their interpretational pitfalls,” *Front. Syst. Neurosci.* **9**, 175 (2016).
116. A. H. Williams and S. W. Linderman, “Statistical neuroscience in the single trial limit,” *Curr. Opin. Neurobiol.* **70**, 193–205 (2021).

117. J. Sun, Z. Li, and S. Tong, “Inferring functional neural connectivity with phase synchronization analysis: a review of methodology,” *Comput. Math. Methods Med.* **2012**, 239210 (2012).
118. S. Ganguli and H. Sompolinsky, “Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis,” *Annu. Rev. Neurosci.* **35**, 485–508 (2012).
119. H. Zeng and J. R. Sanes, “Neuronal cell-type classification: challenges, opportunities and the path forward,” *Nat. Rev. Neurosci.* **18**(9), 530–546 (2017).
120. J. Hirokawa et al., “Frontal cortex neuron types categorically encode single decision variables,” *Nature* **576**(7787), 446 (2019).
121. T. Baden et al., “The functional diversity of retinal ganglion cells in the mouse,” *Nature* **529**(7586), 345–350 (2016).
122. W. Brendel, R. Romo, and C. K. Machens, “Demixed principal component analysis,” in *Proc. of the 24th Int. Conf. on Neural Inf. Process. Syst.*, Curran Associates Inc., Granada, Spain, pp. 2654–2662 (2011).
123. A. H. Williams et al., “Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis,” *Neuron* **98**(6), 1099–1115.e8 (2018).
124. S. Nigam et al., “Rich-club organization in effective connectivity among cortical neurons,” *J. Neurosci.* **36**(3), 670–684 (2016).
125. L. Barnett and A. K. Seth, “The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference,” *J. Neurosci. Methods* **223**, 50–68 (2014).
126. A. Arnold, Y. Liu, and N. Abe, “Temporal causal modeling with graphical granger methods,” in *13th Int. Conf. Knowl. Discov. Data Mining*, p. 66 (2007).
127. L. Barnett, A. Barrett, and A. Seth, “Granger causality and transfer entropy are equivalent for gaussian variables,” *Phys. Rev. Lett.* **103**, 238701 (2009).
128. A. M. Packer et al., “Two-photon optogenetics of dendritic spines and neural circuits,” *Nat. Methods* **9**, 1202–1205 (2012).
129. L. Aitchison et al., “Model-based Bayesian inference of neural activity and connectivity from all-optical interrogation of a neural circuit,” in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, Curran Associates Inc., Long Beach, California, pp. 3489–3498 (2017).
130. Y. Fregnac, “Big data and the industrialization of neuroscience: a safe roadmap for understanding the brain?” *Science* **358**(6362), 470–477 (2017).
131. E. Jonas and K. P. Kording, “Could a neuroscientist understand a microprocessor?” *Plos Comput. Biol.* **13**(1), e1005268 (2017).
132. D. Marr, *Vision*, MIT Press, Cambridge, Massachusetts (1982).

**Yeyi Cai** is currently a PhD student in the Department of Automation at Tsinghua University. Her research employs image enhancement and analysis methods on fluorescence neural imaging and neural decoding. She received her BE degree in control science and engineering from Tsinghua University.

**Jiamin Wu** is currently an assistant professor in the Department of Automation at Tsinghua University. He received his BS and PhD degrees from the Department of Automation at Tsinghua University. His current research interests focus on computational microscopy and optical computing, with a particular emphasis on developing computation-based optical setups for observing large-scale dynamics *in vivo*. He has published over 30 peer-reviewed papers in *Cell*, *Nature Photonics*, *Nature Methods*, *Nature Machine Intelligence*, and so on.

**Qionghai Dai** is currently a professor in the Department of Automation, director of the School of Information Science and Technology, and director of the Institute of Brain and Cognitive Sciences at Tsinghua University. He is also chairman of the Chinese Association for Artificial Intelligence. His research interests include on the interdisciplinary study of brain engineering and the next-generation artificial intelligence. He has developed various multiscale, multidimensional computational imaging instruments and large-scale data analysis methods.