Review of data processing of functional optical microscopy for neuroscience

Hadas Benisty[a], Alexander Song,[b] Gal Mishne[c] and Adam S. Charles[d,*]

[a]Yale Neuroscience, New Haven, Connecticut, United States

[b]Max Planck Institute for Intelligent Systems, Stuttgart, Germany

[c]UC San Diego, Halıcığlu Data Science Institute, Department of Electrical and Computer Engineering and the Neurosciences Graduate Program, La Jolla, California, United States

[d]Johns Hopkins University, Kavli Neuroscience Discovery Institute, Center for Imaging Science, Department of Biomedical Engineering, Department of Neuroscience, and Mathematical Institute for Data Science, Baltimore, Maryland, United States

Abstract. Functional optical imaging in neuroscience is rapidly growing with the development of optical systems and fluorescence indicators. To realize the potential of these massive spatio-temporal datasets for relating neuronal activity to behavior and stimuli and uncovering local circuits in the brain, accurate automated processing is increasingly essential. We cover recent computational developments in the full data processing pipeline of functional optical microscopy for neuroscience data and discuss ongoing and emerging challenges. © *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.9.4.041402]

1 Introduction

Modern neuroscience has been propelled forward by the development of new technologies that offer unique windows into the brain's activity. These new techniques, including high-density electrodes,[1,2] functional ultrasound imaging,[3–6] high magnetic field functional Magnetic Resonance Imaging (fMRI),[7–9] and optical imaging[10,11] all stand to further our fundamental understanding of the brain and represent potential new tools for future therapies. One of the fastest growing directions in this push for advanced neural recording technologies is functional optical microscopy, in particular, *in-vivo* recordings using fluorescence microscopy.

Fluorescence microscopy holds a number of unique advantages over other imaging methods. Unlike fMRI and functional ultrasound, the activity measured is more directly related to neural activity and not as spatially and temporally blurred by the hemodynamic response.[12,13] Furthermore, optical methods do not require inserting probes into the brain, making them typically less invasive than electrophysiology. The drawback to this advantage is the limited penetration depth of the photons. Imaging deeper structures requires more invasive methods, such as implanting a gradient refractive index (GRIN) lens.[14,15] Additionally, fluorescence microscopy provides entire images that capture not just the neural activity, but also the morphology of the cells. Thus fields of view can in theory be registered across days to enable chronic long-term recordings of the same identified neurons, e.g., during learning. However, these high-dimensional and spatiotemporally rich data require significant computational power to extract the core information contained within. Moreover, fluorescence microscopy is actively growing and new volumetric imaging techniques[10,11,16–18] promise to even further increase the scale of such data and the spatiotemporal statistics that must be leveraged in the analysis. To solve this "big data" explosion and to process the ever-growing datasets, there is an ongoing need to meet

---

the challenge of designing robust automated algorithms to accurately extract information from these rich data.

In functional optical imaging of neurons, a fluorescing indicator (typically a protein) sensitive to a biomarker of neural activity is introduced to a cell. Examples of biomarkers include voltage, calcium, potassium, glutamate, sodium, etc.[19–23] Out of these, calcium indicators have been the most widespread. When the level of given biomarker changes, i.e., during or right after a neural firing event, a fluorescent property such as the brightness or emission color of the indicator changes as well. At each image frame, the tissue is illuminated with light at a specific wavelength, and any fluorescing indicators have a probability of raising their energy level. When the energy level falls back to the lower-energy state, light at a longer wavelength is emitted and collected by the microscope. The measured value of the collected light thus reflects the value of the biomarker at a given location.

Practically, the tissue can be illuminated in a number of ways. For example, in single photon widefield microscopy, an entire plane is illuminated at once, using a camera to collect full frames simultaneously. Widefield microscopy can thus acquire high-resolution videos at kilohertz framerates; however, it is limited in its ability to image deeper regions in highly scattering tissues. Specifically, scattering of light in brain tissue greatly blurs images, requiring optical methods that better localize fluorescence at depth. Multiphoton imaging, e.g., two- and three-photon microscopy can penetrate deeper tissue 400 to 1000 μm, but rely on raster scanning technologies to sequentially measure small volumes throughout the plane of imaging.[24–26] Hence, multiphoton imaging is often used to image smaller structures, such as axons, dendrites, and somas, whereas one-photon imaging is used at meso- and cortex-wide scales[24,27–29] or at somatic resolution under challenging optical conditions, such as via endoscopes and miniscopes in freely moving animals.[30,31]

Here, we discuss the emerging line of work that has focused on the task of building the analysis tools required to realize the full potential of high-resolution large scale imaging. The primary goal is extracting time-courses of all the individual units (e.g., neurons, dendrites, brain regions, etc.) from the data so as to relate these to behavior and stimuli in downstream analyses. This is often accomplished by decomposing the movies into two sets of variables: the spatial profiles, representing the area in the field-of-view (FOV) that a unit occupies, and the corresponding temporal fluorescence traces. While this goal is simply stated, the unique properties of *in-vivo* fluorescence imaging create a number of challenges in misalignment of data, tissue aberrations and signal distortion, imaging-dependent noise levels, and severe lack of large-scale ground truth data for validation. These challenges have led to a myriad of approaches, from solving one specific step in the process, e.g., denoising,[32] to whole pipeline implementations.[33] Approaches also range from imaging of specific structures, e.g., widefield[34] to methods aimed at many imaging classes.[35] We aim to provide here a walkthrough of the basic challenges, the landscape of current approaches, and finally emerging challenges with no current solutions.

2 Functional Fluorescence Microscopy

Functional fluorescence microscopy in neuroscience is used to capture the dynamics of neural activity in a wide variety of animal models and targets. Improvements in bioengineering have allowed researchers to study targets from nanometer-sized targets in individually labeled neurons to spanning multiple brain regions across centimeter-wide fields of view. Recordings are taken at several different timescales and framerates, up to hours long recordings at kilohertz rates. Accordingly, specialized microscopes have been developed to tackle these individual experimental requirements.

Most microscopes can be described by two properties: illumination and sampling, of which either one or both may be time-varying. The illumination is a structured light source that aims to excite fluorophores (indicators) in the targeted region of interest without needlessly exciting other areas. The sampling is a mapping of the light emitted from targeted regions onto the sensor. The two most common configurations used for functional imaging today are two-photon laser scanning microscopy (TPM)[36] [Figs. 1(a) and 1(b)] and widefield microscopy [Figs. 1(c) and 1(d)].
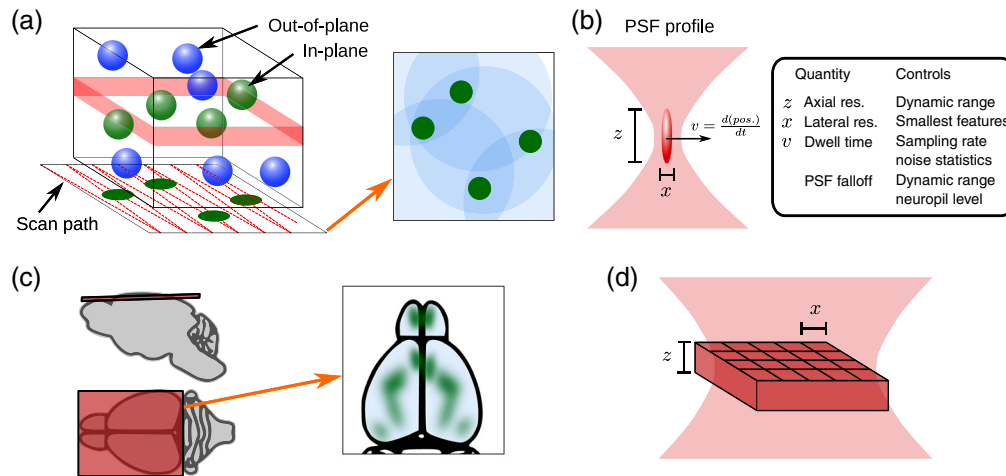
Fig. 1 Optical imaging basics. (a) Laser scanning microscopy (e.g., two-photon microscopy) focuses the laser on a point inside the tissue, sequentially illuminating a chosen plane. Neurons intersecting the plane show up in the rendered images, whereas out-of-plane neurons accumulate as background, and combined nonsomatic activity (e.g., smaller dendritic and axonal components) as neuropil. (b) Choices in the optical illumination impact image properties. The image resolution is defined by the PSF, which is determined by microscope optics and optical properties of the sample. The dwell time controls the time spent at each location and is related to the scanning speed, i.e., the framerate, and can influence the number of photons collected per pixel and thus the signal strength. In laser scanning microscopy, these, in combination with scanning and laser technology, the image size, resolution, framerate, and noise characteristics of the image are determined. (c) Widefield imaging simultaneously illuminates the entire cortical surface plane, capturing dynamics across multiple brain areas. (d) In widefield imaging, camera characteristics primarily determine the image acquisition size and rate. The optical resolution for widefield imaging is strongly limited by microscope optics and optical aberrations from the sample.

In TPM, the illumination is a focused, pulsed near-infrared (NIR) laser that is raster-scanned in a two-dimensional (2D) pattern [Fig. 1(a)] and the sensor is a single pixel detector (typically a photo-multiplier tube) collecting all emitted light from the whole FOV. The fluorophores are excited via two-photon absorption,[37] which has a squared-law relationship with the intensity of the excitation beam. This results in the absorption process being confined to a small, point-like focal volume [Fig. 1(b)]. As emitted light is typically bulk-collected in TPM, the excitation focal volume describes the point-spread function (PSF) for these systems. TPM has been widely adopted primarily due to its excellent optical sectioning properties and use of NIR wavelengths. Because scattering of NIR light is much lower in brain tissue than visible light and scattering of the fluorescence does not impact image quality, TPM achieves a practical imaging depth of up to 500 μm, even in the highly scattering mammalian or avian brain.

In widefield microscopy [Fig. 1(c)], the illumination is a visible LED or laser that excites the imaged volume and the sensor is a camera that maps onto the target volume. The excitation light propagates through the whole imaging volume, strongest near the surface of the sample, and is attenuated over a couple of 100 μm into the sample via tissue absorption. The emitted light is imaged onto the camera, with the contribution from the focal plane mapped most closely onto the sensor and out-of-focus light diffusely collected as well [Fig. 1(d)]. The PSF for the system is determined by the imaging optics, with light from the focal plane having the sharpest response and blurring away from the focus. The image quality is further degraded by scattering in the emission. In reasonably transparent animal models (such as *C. elegans* or larval zebrafish), this effect is minimal. In highly scattering tissues, like the mammalian brain, this effect limits the overall depth with acceptable imaging quality for neuron-scale recordings to typically 200 μm. While the emitted light frequently comes from the whole volume, as in densely labeled samples, genetic targeting or microscopy techniques can isolate the signal to a smaller population of cells. A major advantage of widefield microscopy is the ease of setup in a variety of animal models and scalability to very large areas and framerates.

Several microscopes have been designed to optimize functional neural recordings. For TPM, development has focused on improving the acquisition speed and volume size. The illumination shape of the PSF has been transformed for imaging sparse tissues effectively[38] or for volumetric imaging.[16] Three-photon microscopy enables even deeper imaging.[39] Changing the illumination time course with custom scanning paths can make fast jumps from cell to cell.[40] Live updating of both the scanning pattern and timecourse is used for high-speed image acquisition.[41] For collection, cameras have been used to replace the single-pixel detectors to improve acquisition rates.[42] Recently, the combination of multiple techniques and technologies has reported cellular resolution imaging of up to a million neurons simultaneously.[11]

Changes to the techniques of widefield microscopy have focused on optical sectioning and volume scanning. Optical sectioning can be achieved by structuring the illumination such that only the plane mapping onto the camera is excited. This has been achieved primarily through light-sheet microscopy, which has been scanned to image three-dimensional (3D) volumes in larval zebrafish[43] and also with a single objective in mice.[10] Light field microscopy has been used to structure mapping of the sampling from the imaged region to the camera[44] and in combination with confocal microscopy enables high-quality optical sectioning.[45] One of the most major developments has been the miniaturization of microscopes (miniscopes) for head-mounted functional imaging, which has greatly expanded the types of problems and brain areas researchers can explore.[46–48]

All of the advances outlined have enriched the space of acquired brain recordings, especially when intersected with the range of available indicators. Specifically, they create a range of spatiotemporal resolutions, signal quality, and imaged morphology, a subsection of which
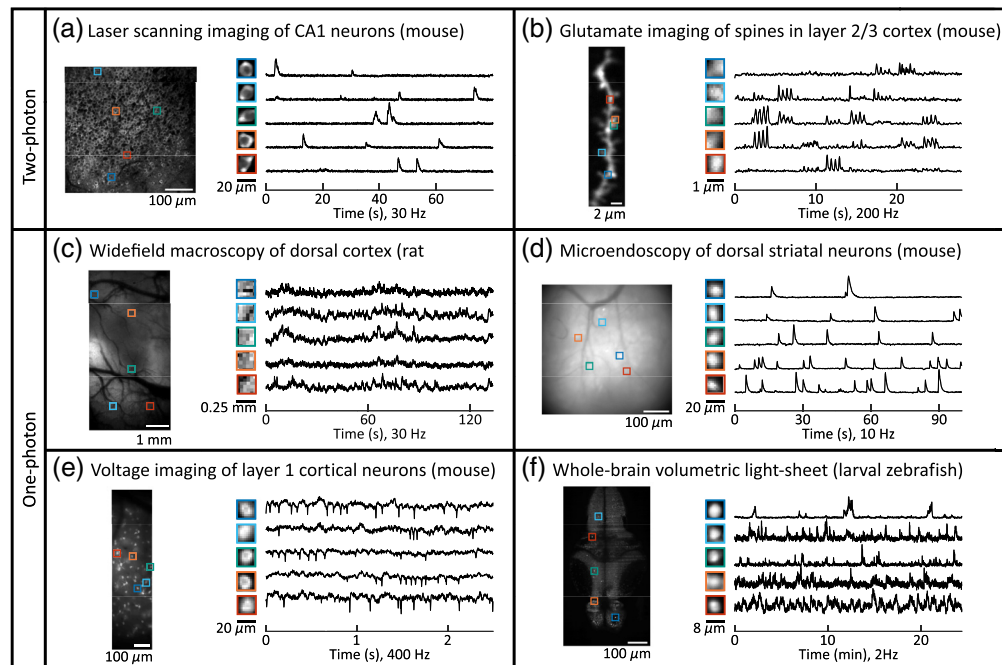
Fig. 2 Examples of functional microscopy data. Functional microscopy data obtained via different modalities, at different scales, and in different animal models expressing different indicators can take on a diversity of signal characteristics. Shown are six examples spanning a range of datasets. (a) Laser scanning two-photon calcium imaging (GCaMP5) in mouse hippocampus.[49] (b) Glutamate (iGluSnFR3.v857) spine imaging using laser-scanning two-photon microscopy in layer two-third mouse cortex.[50] (c) One-photon widefield calcium (GCaMP6f) imaging of rat dorsal cortex using a head-mounted macroscope.[48] (d) One-photon calcium imaging using a head-mounted microendoscope in mouse striatum.[51] (e) One-photon voltage imaging (Voltron525-ST) in layer 1 of mouse cortex.[52] (f) Whole-brain light-sheet calcium imaging (GCaMP6f) in larval zebrafish.[53] Only one of 29 imaging planes each recorded at 2 Hz is depicted here. Note that the spatial and temporal statistics range dramatically across domains, requiring flexible data processing approaches.

we depict in Fig. 2. To bridge the gap between the raw data and scientific discovery, these data must be analyzed to extract useful neuronal activity.

3 Fluorescence Microscopy in Space and Time

The current state of functional fluorescence microscopy analysis can be understood through the long history of developments that have led to today's state-of-the-art. Initially, fluorescence microscopy had long scan times and was thus focused on imaging static tissue (see Ref. 54 for a review of the very early history of fluorescence microscopy for biology). Thus, all the relevant information was anatomical in nature and could be traced manually or, later on, identified automatically to discover complex biological structures. Fluorescence microscopy continues to play a vital role in imaging static anatomical structures and morphological fitting algorithms have evolved with the technology.[55–57]

As scan times improved and the ability to inject fluorescent indicators into live tissue emerged, fluorescence microscopy expanded to the imaging of time-varying biomarkers, e.g., voltage indicators in the 1970s[58,59] and calcium indicators in the 1980s.[60] With this shift came the added dimension of time, as now fluorescence was a temporal quantity. In neuroscience, this enabled the study of the activity of the neurons over time. Initially, however, labeling methods were in their nascent stage and the use of viruses to introduce fluorescing proteins into cells created a high variability in labeling strength and longevity. The result was often sparsely labeled tissue where overlapping neurons were rare. Without overlap, isolating a single cell's activity could thus be accomplished by identifying individual neurons anatomical extent in the video (e.g., based on the temporal mean or variance of the imaging stack), and then averaging the pixels belonging to each neuron to extract their time-trace.

To see this, we consider the statistical assumption that each neuron is modulated in a linear way (i.e., the spatial shape and temporal fluctuations are independent). Each N-pixel frame $y_t \in \mathbb{R}^N$ at time t, stacked as an $N \times 1$ vector, can be thought of as the linear combination of the d cell shapes $a_i \in \mathbb{R}^N$ for $i = 1, \ldots, d$ and its activation ϕ_{it} at time t:

$$y_t \asymp \sum_{i=1}^{d} a_i^T \phi_{it} = A\boldsymbol{\phi}_t, \tag{1}$$

where $A \in \mathbb{R}^{N \times d}$ is a matrix consisted of the spatial components and $\boldsymbol{\phi}_t \in \mathbb{R}^N$ is the activation of all cells at time t. Solving a least-squares optimization of the form

$$\hat{\boldsymbol{\phi}}_t = \arg \min_{\boldsymbol{\phi}} \|y_t - A\boldsymbol{\phi}\|_2^2 \tag{2}$$

for the activity of all cells $\hat{\boldsymbol{\phi}}_t$ at time t can then be written as the pseudo-inverse $\hat{\boldsymbol{\phi}}_t = (A^T A)^{-1} A^T y_t$. With no overlap the Gram matrix, $A^T A$ is simply a diagonal matrix with the norms of the cell shapes on the diagonal. Thus, the activity for the ith cell at time t is simply a weighted average of the pixel values at that time-point, which has been exactly the methodology with hand-drawn spatial profiles, sometimes also called regions of interest (ROIs).

As labeling methods advanced and indicator designs provided brighter fluorescent tags, overlaps between neurons and with other processes (e.g., dendrites) became more common. The increased cell count vastly increased the burden of manual annotation and the increased overlaps mathematically results in a nondiagonal Gram matrix $A^T A$, removing the validity of direct averaging. Thus, new conceptual approaches were required for cell identification. Rather than solely relying on spatial cues to isolate neurons, activity had to be demixed in space *and* time. In fact, as opposed to purely anatomical studies, many modern systems neuroscience analyses abstract away from space, analyzing a neuron-×-time matrix irrelevant of anatomy. Methods such as dimensionality reduction,[61–63] dynamical systems analysis,[64] statistical connectivity,[65] etc. operate on the time-traces and thus suffer more from inexact estimation of neural activity rather than inexact estimation of spatial shapes.

Under the statistical assumption that each neuron is modulated in a linear way, i.e., the spatial profile is constant over time with brightness controlled by a time-vector representing the cell's activity, the problem of isolating neurons can be considered as a matrix factorization problem. If the data matrix $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T] \in \mathbb{R}^{N \times T}$ is the pixel-by-time fluorescence video matrix, each neuron contributes one rank-1 component

$$\boldsymbol{Y} \approx \sum_{i=1}^{d} \boldsymbol{a}_i \boldsymbol{\phi}_i^T = \boldsymbol{A}\Phi^T, \tag{3}$$

where $\boldsymbol{a}_i \in \mathbb{R}^N$ is again the i'th neuron's spatial profile (how it appears visually in the data), and $\Phi = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_N] \in \mathbb{R}^{N \times T}$ is the matrix where the i'th column, $\boldsymbol{\phi}_i \in \mathbb{R}^T$ is the i'th neuron's time-trace (how the biomarker-driven fluorescence changes over time). To find the rank-d decomposition, Mukamel et al.[66] used a combination of PCA and ICA: first, PCA was used to reduce the overall dimension of the data and to obtain an initial guess $\boldsymbol{Y} \approx \sum_{i=1}^{d} \boldsymbol{u}_i \boldsymbol{v}_i^T$, where $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the left and right singular vectors, respectively. ICA was then performed on the set of vectors $\boldsymbol{z}_i = [\alpha \boldsymbol{u}_i^T, (1-\alpha)\boldsymbol{v}_i^T]^T$, i.e., identifying a rotation of $\boldsymbol{Z} = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_d]$ that makes the columns more independent, and the parameter α trades off importance of the spatial and temporal components obtained via PCA (i.e., the left and right singular vectors). The rotation in the ICA step transforms the PCA components into their independent components, demixing both the spatial and the temporal components at once. This procedure allowed for overlapping components and included temporal independence when identifying neurons.

The matrix factorization approach continued to expand,[67–69] primarily formulated as an optimization program:

$$\{\hat{\boldsymbol{A}}, \hat{\Phi}\} = \arg \min_{A,\Phi} \|\boldsymbol{Y} - \boldsymbol{A}\Phi^T\|_F^2 + \mathcal{R}_A(\boldsymbol{A}) + \mathcal{R}_\Phi(\Phi), \tag{4}$$

where $\mathcal{R}_A(\boldsymbol{A})$ and $\mathcal{R}_\Phi(\Phi)$ represent appropriate regularization terms over space and time, respectively, that can vary between methods, and often include terms such as the component norms, number of components, sparsity, non-negativity, and spatial cohesion. As direct optimization is often difficult for problems of this size, alternating descent type algorithms are often employed, i.e., iteratively solving

$$\hat{\boldsymbol{A}}|\hat{\Phi} = \arg \min_{A} \|\boldsymbol{Y} - \boldsymbol{A}\hat{\Phi}^T\|_F^2 + \mathcal{R}_A(\boldsymbol{A}),$$

$$\hat{\Phi}|\hat{\boldsymbol{A}} = \arg \min_{\Phi} \|\boldsymbol{Y} - \hat{\boldsymbol{A}}\Phi^T\|_F^2 + \mathcal{R}_\Phi(\Phi).$$

Positive aspects of this approach include that each of these optimization problems can be solved reasonably efficiently (conditioned on judicious choices of regularization) and that different assumptions on the time-traces and component shapes can be incorporated naturally via regularization.

Thus, in the current landscape of algorithms that extract neural signals from fluorescence microscopy data we now see a combination of approaches. One approach focuses primarily on anatomical identification, leaving the identification of functional traces as a later stage, and another which places identification of the time-traces on equal ground (or more) with the spatial maps and extracts the two jointly.

3.1 Methods Focusing on Space

The class of methods that focus on anatomical identification have been mostly inspired by image segmentation, using both classical and modern approaches. Each of these methods relies on a spatial model of the data, either preset or learned from the data.

One such approach, dictionary learning, stems from the broader image processing literature.[70] Dictionary learning assumes a sparse generative model for image patches[71] and has been applied to calcium imaging to learn spatial dictionaries whose atoms represent neuronal shapes. The identified neuron shapes can then be used to estimate corresponding time courses.[72–74]

An early approach ignored temporal information entirely and developed a spatial generative model based on convolutional sparse block coding,[72] applied to normalized mean projections of the imaging data. A dictionary of cellular shapes was learned with corresponding locations of these shapes in the mean image.

Modern deep learning methods for image segmentation have also been adapted to the ROI extraction problem. These are supervised approaches that require labeled training data. UNet2DS[75] is a neural network based on the fully convolutional UNet[76] model, an architecture that was developed for biomedical image segmentation. UNet2DS takes the mean image as an input and outputs two probability masks of a pixel belonging to either a cell or the background. Since U-Net2DS ignores temporal information, it has difficulty separating overlapping neurons and detecting sparsely firing cells. Such methods also cannot differentiate active cells, i.e., cells that exhibit deviations from baseline fluorescence consistent with spiking events, from non-active cells.

3.2 Methods Using Time Information to Identify Spatial Masks

Another class of methods still emphasizes the identification of the spatial components, however, they rely on temporal information, e.g., in calculating correlations between pixels in the FOV based on their temporal activity.

One approach is to view cell detection as a clustering problem.[77,78] For example, in HNCcorr[78] after selecting a set of seeds (superpixels), this method aims to find a cluster within a patch containing the seed by solving the Hochbaum's normalized cut (HNC) problem, which is similar to normalized cut. The HNC formulation balances a coherence term, which maximizes the total similarity of the pixels within each cluster, with a distinct term that minimizes the similarity between the cluster and all others, where similarity is based on time-trace activity. Local selective spectral clustering (LSSC)[79] also solves the cell identification problem as clustering pixels in a high-dimensional feature space. Following the construction of a sparse graph whose nodes are the pixels and whose weights are determined by pairwise similarity of the time-traces of the pixels, the nodes are embedded using the eigenvectors of the random-walk graph Laplacian. Pixels are then clustered together in an iterative approach based on selecting subsets of the eigenvectors that best separate a cluster from the rest of the data, and enabling overlapping clusters. After detecting neuronal components (the graph construction allows for morphology-agnostic clustering), time traces are demixed and calculated by projecting to a low-rank space.

Activity-based level set segmentation (ABLE)[80] is an image segmentation approach relying on active contours.[81] ABLE defines multiple coupled active contours in the FOV where an active contour seeks to partition a local region into an interior corresponding to a neuronal component and a local exterior, such that the pixels within each region are similar to one another temporally. The evolution of the contours is performed by the level set method,[82] where only neighboring cells directly affect a cell's evolution. ABLE handles overlapping cells by coupling the evolution of active contours that are close to one another. ABLE automatically merges two cells if they are close and temporally correlated and prunes cells if their size is too small or too large. An advantage of this method is that it makes no assumption on a cell's morphology or temporal activity, therefore it can potentially generalize to different indicators and spatial morphologies.

Convolutional deep learning networks taking into account temporal statistics or activity have also been developed in recent years.[83–85] A (2+1)D convolutional neural network[83] was trained on spatiotemporal sliding windows and the output represented the probability of a pixel belonging to an ROI centroid. Apthorpe et al.[83] demonstrated that adding the temporal domain helps suppress noisy detections compared with a 2D network that only took as input the time-averaged image. However, the network they trained only learned spatial 2D kernels.

In comparison, STNeuroNet[84] is a 3D convolutional neural network based on DenseVNet (similar to a U-Net) also trained on overlapping spatiotemporal blocks, and its output is a binarized probability map. By adding a temporal max-pooling layer to a typical DenseVNet architecture, the spatiotemporal features are reduced to a spatial output, which increases the speed of training and inference. This gain is important for network validation and low-latency inference in closed-loop experiments. Shallow U-net neuron segmentation (SUNS)[85] aims to simplify the neural network architecture of STNeuroNet to further improve speed while still incorporating

temporal information. To this end, a temporal matched filter tailored for shot noise is applied to the input video, thereby enhancing calcium transients occurring across multiple frames, and reducing temporal information into individual 2D frames. Preprocessing also includes whitening, which normalizes the fluorescence time series of each pixel by the estimated noise of that pixel. This yields an SNR representation that highlighted active neurons and obscured inactive neurons. The SNR video is then processed by a shallow UNet, whose output is a probability map.

Postprocessing for both of these deep-learning methods includes clustering of the output probability maps to detect individual cells and aggregation of detected cells across the probability maps for merging and pruning. Finally, deep learning, instance segmentation, and correlations (DISCo)[86] uses a combination of time-trace correlation-based pixel segmentation on a graph and a convolutional neural network to identify individual spatial profiles. Note that while these methods identify cells based on a spatiotemporal analysis, they do not address the issue of estimating time-traces.

3.3 Methods Focusing on Space and Time

Rather than leave time-trace estimation as a secondary step, spatiotemporal demixing methods aim to simultaneously identify the spatial maps with their corresponding fluorescence traces.

One such method builds off of the dictionary learning framework first used in spatially-focused cell identification. Specifically, Diego et al.[73] generate a space-time dictionary learning extension of convolutional sparse coding to video data with nonuniform and temporally varying background components. Petersen et al.[74] approach the dictionary learning of spatial components via iterative merging and clustering. The former of these extracts sparse spike trains in the convolutional dictionary. The latter mainly emphasizes identifying the spatial maps but, as a by-product of the decomposition algorithm, computes the time traces as well.

Many algorithms that jointly estimate both spatial profiles and time traces aim to optimize the cost of the form of the optimization program in Eq. (4). In particular, many of these methods follow developments in analyzing somatic calcium imaging data, which are based on matrix factorization with non-negative constraints that represent the expected positive deviations from baseline: non-negative matrix factorization (NMF).[35,67–69,87–89] The main difference between these methods is the specific choices of regularization and the specific implementation of the algorithmic steps. For example, multiple methods follow the aforementioned alternating optimization,[35,67,69,87] whereas another approach instead uses semidefinite programming.[68] Versions of matrix factorization methods have been perhaps the most widespread, with special versions being designed for one-photon endoscopy data,[90] widefield data,[34,35] and voltage data.[91]

Regularization has been applied both on the temporal and spatial dimensions. Temporally, sparsity over the time traces has been a popular constraint. Specifically, neurons often fire infrequently, relative to the number of frames in a video.[16,67] Taking advantage of this sparsity requires accounting for the biophysically-induced exponential decay of fluorescence. While more sophisticated methods have been developed for postprocessing deconvolution (see Sec. 4.5), a simple way that has been applied to cheaply deconvolve data is based on the observation that exponentially decaying responses (i.e., a single-pole response function) with decay rate α can be expressed as the difference equation $F(t) = \alpha F(t-1) + S(t)$, where $S(t)$ is the instantaneous impulses over time. Thus simply computing $\tilde{S}(t) = F(t) - \alpha F(t-1)$ should give a reasonable estimate of a sparse impulse train. The parameter α can further be tuned to maximize sparsity $\max \|S(t)\|_0$.

Spatial regularization is more complex and often requires a model for neural processes of interest. Initial ideas included ensuring connected components, or by minimizing the total variation (TV) norms to minimize the complexity of shapes and remove spurious pixels.[68] In more recent versions of matrix factorization algorithms, deep networks trained on human-annotated labeled data have been used to classify true and false cells,[49] the resulting system serving as a posthoc filter that effectively double-checks the spatial regularization.

As an alternative to shape-based regularization, recent work has leveraged sparsity and spatial correlations in a new way.[35] Instead of using spatial locations of pixels to constrain the components' spatial profiles, graph-filtered temporal (GraFT) dictionary learning builds a graph

among all pixels that enables like-pixels to share time-trace decompositions. This approach effectively gives up completely on space and instead focuses on the learning of a dictionary of time traces over the data-driven pixel graph.

More recently, another important modification to the base optimization in (4) has been emerging. Rather than focusing on changing the regularization terms used, the least-squares form of the data fidelity term has instead been reconsidered. The least-squares cost enforces a linear-Gaussian data generation hypothesis, however, a number of nonlinearities in fluorescence dynamics, incompleteness in component discovery, and the non-Gaussian statistics of the photo-diodes all contribute to various extents of errors in demixing. Four main alternatives have emerged, including a robust-statistical approach leveraging a Huber cost function,[87] a contamination-aware generative model approach,[92] a zero-Gamma mixture model,[93] and a deep-learning approach.[94]

4 Imaging Analysis Pipeline

The fundamental output of an optical functional imaging analysis pipeline is the identified spatial profiles and more importantly their corresponding time traces. While demixing thus serves as the core of analysis, often multiple preprocessing and postprocessing steps are typically part of the pipeline to facilitate this output (Fig. 3). Motion correction is typically the first step in the analysis pipeline, to register all frames in the imaging stack such that the neuronal components to be extracted occupy the same spatial footprint in all frames. Denoising can be applied as a preprocessing step either temporally[35] or spatially[79] to improve the detection of ROIs. Normalizing per-pixel time-traces, e.g., by z-scoring, can enhance dim cells, and improve cell detection. Following ROI extraction, postprocessing in the temporal domain can include the following: neuropil estimation, denoising of time traces (if not implicitly part of the extraction itself), and calculating normalized time traces ($\Delta F/F$). Postprocessing in the spatial domain can include automatic classification[49] of identified components into true or false neuronal components or manual quality control. To identify single spiking events, deconvolution is a postprocessing step that aims to recover sparse firing events from the fluorescence time traces. Finally, in longitudinal studies, postprocessing in the spatial domain can include registration of imaging sessions and matching of ROIs across sessions.

4.1 *Motion Correction*

The majority of ROI extraction methods rely on neuronal components being within a fixed/consistent spatial footprint in the FOV – thus necessitating image registration of the individual frames. Motion between frames can be due to several factors:[95–99] animal motion during imaging (e.g., locomotion), scanning artifacts, mechanical strain, drift relative to the objective, changes in the brain, e.g., due to hydration, etc. Given the length of imaging sessions (tens of thousands of time frames), computational complexity of the registration approach is an important consideration. For fast acquisition rates, interframe motion can be considered as a global constant offset of the FOV, therefore rigid registration of translational shifts is sufficient.[100–102] For example, registration can be performed to a reference (template) image using cross-correlation or phase-correlation.[69] The reference image is typically set as an average over the initial frames and then regularly updated as an average over later subsets of frames, or an average over the full stack. The reference frame can be made more precise by an iterative refinement procedure to reduce blurring.[69,103] As an alternative to performing correlation-based translation registration, bright cells can be detected and tracked over time using particle tracking.[104] A rigid (translation and rotation) transform is then calculated for each frame to the next by minimizing the residual displacements of all tracked cells. This approach can also be applied to volumetric imaging.

Nonrigid motion artifacts can occur due to lower acquisition speeds, which result in artifacts such as shearing in later scanned lines of the image, or slow distortions over long recording sessions due to mechanical (z-drift with respect to the objective) and biological issues (warping of brain tissue due to metabolic activity, dilation of blood vessels, and liquid reward consumption).[105] Nonrigid motion correction usually relies on splitting the image into overlapping spatial
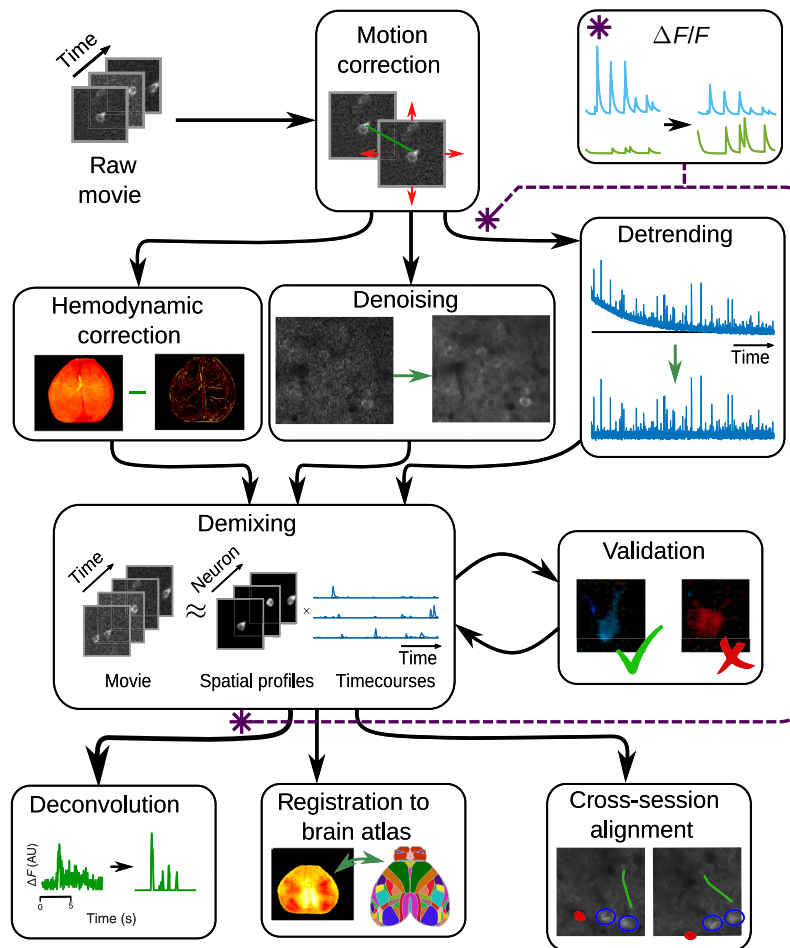
Fig. 3 The optical imaging pipeline, expanded. The first step in analyzing fluorescence microscopy data is to correct any motion, reducing interframe misalignment. Between motion correction and demixing (i.e., cell-finding), steps to remove imaging noise and confounding factors are taken. Specific steps include de-trending to remove photobleaching, denoising to increase SNR, and hemodynamic correction in widefield data. The core pipeline stage of focus is demixing or identification of individual fluorescing components. Demixing should always be paired with some sort of validation on the output to prevent artifacts and other errors. The output of the demixing can be used to infer firing events (in single-cell resolution imaging), align multiple imaging sessions, or register to a global brain atlas (in widefield data). One important step: the $\Delta F/F$ calculation, which normalized fluorescence traces to their baseline, can be computed before or after demixing.

patches and performing registration at the patch level. This registration can be rigid at subpixel resolution[69,106] or a more flexible affine transform.[103]

Most methods for motion correction target two-photon somatic imaging. However, such methods can struggle with nonsomatic neuronal components such as dendrites, due to the difference in size and impact of z-drift. While a cell-body is typically on the order of $\approx 15\ \mu$m—with variation depending on brain area, species, etc.—the width of an axon or a dendrite is $\approx 1\ \mu$m. Thus, slight registration errors can have a significant effect on identifying the spatial footprint of these components. Furthermore, z-drift can cause segments of tuft dendrites to shift in and out of the FOV. This can lead to difficulties in aligning the dendrite to the reference image. From a computer vision perspective, this can be thought of as image registration under occlusions.

Another potential complication is the use of GRIN lenses to image deeper structures in the brain. Optical aberrations near the edges of GRIN lenses can significantly change the motion characteristics, requiring distortion-aware realignment. Finally, niche technologies can also create novel situations, such as the rotating platform developed for near-freely moving imaging.[107,108]

4.2 *Denoising and Normalization*

The absolute noise levels present in optical imaging create a challenging signal extraction environment. It is only because many sequential frames of the same population are recorded that individual cells can be identified. This process, however, can be improved by modest noise filtering as a preprocessing stage. A number of methods (with example applications) are used across the literature, including median or low-pass filtering,[16,79] downsampling, PCA projection,[69,79] *z*-scoring, wavelet denoising,[35] other hierarchical models,[109] and deep learning-based denoising.[32,110] All these approaches make different noise and signal model assumptions and should be used judiciously.

For example, median/low-pass filtering and downsampling, are simple, quick steps that can be run on the data at the early stages. Median filtering is effective at reducing shot noise common in low-photon environments, at the cost of making shapes in the image more "convex" (i.e., filling in corners). Low-pass filtering and downsampling reduce high-frequency noise, however, blur the data via the convolutional filter. Downsampling[83,111] has the further benefit of reducing the data size in space or time, thus reducing later computational costs, however, effectively reduces the sampling resolution.

Other signal models are more complex. Hierarchical models can provide flexible ways of both incorporating different noise classes (e.g., Poisson) and flexible signal models (e.g., via interpixel correlations), however, at a heavy computational overhead.[109] Wavelet-based denoising[112,113] is perhaps the most versatile in this class, as both per-image and per-pixel time-trace denoising are computationally efficient, readily implemented across programming languages, and can handle sharp transitions, thereby reducing blurring.

Yet other methods model the noise in ways based on the data itself. PCA-based projections identify a low-dimensional space that captures much of the data variance, effectively assuming that low-variance principal components represent noise. Sparsely firing cells or dim cells, however, can often end up in low-variance components and thus removed with the noise. Penalized matrix decomposition[91] denoises the imaging data by performing a patch-wise penalized low-rank decomposition. Recently, general advances from deep learning[32,110,114] have been proposed for denoising calcium imaging.

For example, DeepInterpolation[32] uses a clever design that trains a neural network to predict a movie frame based on previous and past frames. Independent noise, which cannot be predicted, is thus filtered out. A broader image-restoration method CARE,[115] uses a U-net trained on high- and low-resolution images to enhance signal quality. As with all deep-learning methods, the drawbacks include training the network (if the pre-trained options do not fit the application) and a general lack of knowledge as to the exact expected biases of the black box system.

Usually performed at the same time as denoising, normalization plays an important role in the numerical stability of algorithms. Having data with large (or small) overall fluorescence values can cause problems in the condition numbers of the matrices used in optimizing costs in the demixing step. Normalizing, e.g., to unit-median or unit-max values help constrain these effects, and all normalization can be undone after demixing to return meaningful fluorescence values to the identified time-traces. As an example, methods based on deep networks typically employ preprocessing to ensure appropriate dynamic range across the data, regardless of background or other inhomogeneities in the illumination (e.g., via homomorphic filtering[85]).

As a final note, a form of normalization that often is required in preprocessing is detrending to remove photobleaching[116] (e.g., see Refs. 117 and 118). Photobleaching is the reduction in overall fluorescence over time stemming from the fluorescent proteins becoming trapped in an intermediate quantum state and becoming inactive. The result is a large imbalance over time in the dynamic range of the signal, which can hinder most demixing algorithms.

4.3 *Neuropil Estimation*

In 2p imaging, neuropil is a "background" signal that contains the fluorescence of neuronal elements that are out-of-focus (e.g., dendritic and axonal) and scattering. This signal can contaminate the estimated fluorescence of an ROI if it is not properly accounted for. Some methods estimate neuropil as part of the ROI extraction process by automatically identifying the signal of

the exterior surrounding the cell,[80] or explicitly or implicitly adding at least one background signal in the linear decomposition model,[35,49,67,89] e.g.

$$Y \asymp \sum_{i=1}^{d} a_i \phi_i^T + a_b \phi_b^T,\tag{5}$$

where a_b and ϕ_b are the spatial and temporal components of the background, respectively. In other methods, neuropil is estimated in a posthoc process by calculating the average time trace from the pixels within the ring surrounding each extracted spatial profile.[69,119] The signal is then subtracted from the time trace of the spatial profile. Care should be taken in neuropil subtraction not to over-correct. Over-correction can often be identified by noting significant negative dips in the fluorescence baseline

4.4 *Normalized Fluorescence (ΔF/F)*

Neurons have different concentrations of fluorescence indicator, e.g., GCaMP, which can result in varying levels of fluorescence both across cell populations and between individual neurons.[120,121] Additional variation also occurs across the image field of view due to imaging technique, microscope optics, and sample variation.

Therefore, to obtain comparable signals from varying neural sources, extracted time-traces are normalized[122] to remove the baseline fluorescence activity and adjust the amplitude as follows:

$$\frac{\Delta F}{F}(t) = \frac{F(t) - F_0}{F_0},\tag{6}$$

where $F(t)$ is the extracted time trace and F_0 is the baseline fluorescence.

Calculation of baseline fluorescence relies on the assumption that neurons have sufficiently sparse activity, and thus a baseline level that does not correspond to neuronal activity can be estimated using varying heuristics. A common method of estimating F_0 is averaging the activity of the time frames with $p = 10\%$ lowest activity (the exact value of p can vary from lab to lab, for widefield imaging $p = 50\%$ may be used to instead describe the percent change from the mean activity level). This method assumes neurons are quiescent for long enough periods that the change in fluorescence level is effectively zero at this percentile. A lower percentile is not used to account for error caused by sources such as shot noise or axial motion.

The running percentile calculation additionally assumes there is little to no fluorescence from other sources. In datasets that are densely labeled or have high background fluorescence, the technique will consistently underestimate the true $\Delta F/F$ as contributions from other sources get added to the baseline estimate. Estimating the baseline fluorescence with overlapping sources requires separating the fluorescence contributions from individual sources and any background source, which is automatically achieved in most segmentation algorithms.[67,69] This approach may be more reliable than the running percentile calculation by integrating additional information provided by activity transients. A test using synthetic data demonstrated that these estimates of baseline fluorescence were reliable for the brightest cells, and another estimation procedure making use of pixel-wise spatial information further improved estimates of baseline activity.[123]

4.5 *Deconvolution*

The product of the core demixing stage includes a set of time traces—one per component—that contain the temporal fluorescence fluctuations. Due to the complex interactions and resulting buffering of nonvoltage fluorescence-inducing biomarkers, the transient increases in fluorescence due to individual spiking events are stretched out over time. For example, calcium indicators can observe deviations from baseline in the recorded fluorescence from a single spiking even for seconds.

As spiking events are typically considered the primary source of neural communication, efforts to infer underlying spiking from fluorescent data have emerged, taking a number of

forms.[73,124–131] The attempt to invert a generative model of fluorescence $F(t)$ as a function of spike times $\{\tau_1, \tau_2, \ldots, \tau_K\}$ is fundamental. The full biophysical model dictates that at each time τ_k of a spike event, a stochastic influx of the biomarker (e.g., calcium) flows into the cell and drives nonlinear differential equations and nonlinearity that determine the level of bound fluorescent protein over time. While the full biophysical model includes nonlinear differential equations and nonlinearities,[123,132,133] this model can be linearly approximated as

$$F(t) = \sum_{k=1}^{K} g_k h(t - \tau_k) + e(t), \tag{7}$$

where $h(t)$ is the response curve (e.g., exponential rise-and-decay) to a single event and g_k represents the stochasticity of the influx of biomarker at the k'th event.

Given the convolutional form of this model, identifying the time-points of events has taken the form of deconvolution. These algorithms have taken the form of assuming functional forms over $h(\cdot)$, such as exponential[67] or double exponential[124] kernels, or taken more model-free approaches in a deep-learning framework.[130,131,134] Specific algorithms have also varied widely, including exact ℓ_0-regularized optimization,[127] marked point process,[129] interior point optimization,[135] active-set methods,[136] and variational autoencoders.[134]

One of the core difficulties in deconvolution is the probabilistic relationship between spikes and fluorescence. In addition to the biomarker levels being variable, simultaneous recordings of electrophysiology and calcium imaging show that at times optical imaging can miss single events at significant levels (i.e., missing 70% to 80% of individual spikes).[137] Thus, in general across noise, indicators, and other experimental conditions, deconvolution may have highly varying performance limits. In fact, for assessment, benchmarking attempts have avoided direct timing comparisons, opting instead for using local rate averages over >40 ms bins as a validation metric.[130] Interestingly, spike ambiguity in optical imaging has seeded another approach: to remove spiking events from the equation—literally—by marginalizing out the spiking events and directly estimating a latent spike rate.[138]

4.6 Multisession Registration

The ability to record from the same identified population of neurons with functional optical imaging in longitudinal experiments across multiple days is a major advance and enables understanding of long-term processes such as learning and memory. One of the crucial postprocessing steps of such longitudinal experiments is the alignment of the recorded imaging data across days to enable one-to-one mapping of neurons across all sessions. This is essential to understanding the changes in neural representation over time. However, alignment is challenging due to the 3D nonrigid transformations between imaging sessions. These are a result of, for example, day-to-day variance in the imaging angle due to slight changes in the angular placement of the microscope objective, day-to-day variance in optical clarity of cranial windows, and changes in the brain tissue over days. Specifically, in TPM, this can lead to differences in the shape of recorded cells since slight z-drift and tilts can result in relatively large changes in the cross-section of a cell. Semiautomated approaches to calculate the transform between imaging sessions and match neurons exist, however these rely on user input to select matching ROIs and only align pairs of sessions.[69] Recent methods propose registering imaging sessions based on fully affine invariant methods originally developed for natural image registration.[139] A recent approach[140] based on the classical SIFT[141] algorithm enables fast automatic registration of calcium imaging sessions and one-to-one matching of ROIs, even if the neuron was not detected in all sessions.

In one-photon imaging, alignment is challenging due to light scattering and lack of optical sectioning, which increase the similarity between the time traces of neighboring neurons in the FOV.[142] In addition, only active cells can be tracked, as opposed to multiphoton imaging. Sheintuch et al.[142] developed a probabilistic method for automated registration across one-photon imaging sessions that is adaptive and optimized to different datasets. First, all cells are mapped to the same image by registering each session to a reference session using a rigid transformation based on the centroid locations of extracted ROIs. Next, the probability for any pair of neighboring cells from different sessions to be the same cell is calculated, given their spatial

correlation and centroid distance and a probabilistic model for similar and dissimilar matched cells. Cells are finally aligned across sessions by an iterative procedure based on the estimated probabilities.
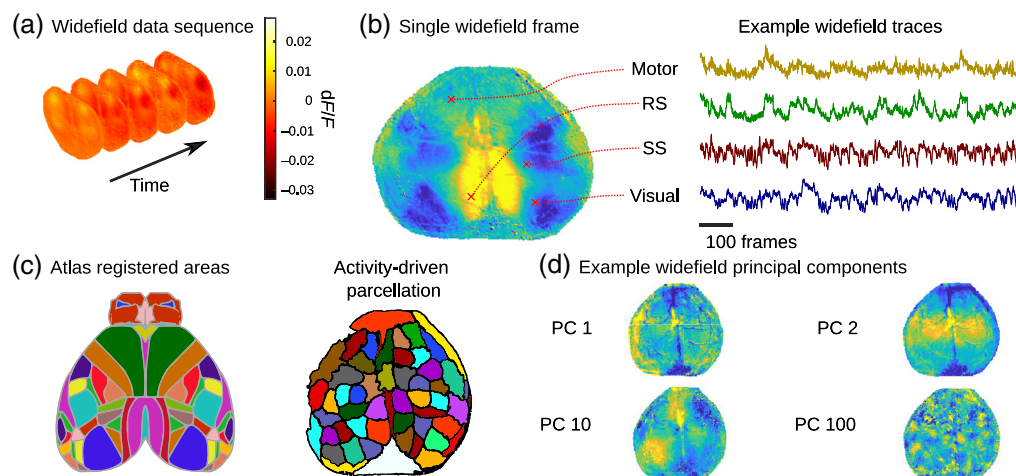
5 Widefield Imaging Analysis

Multiphoton microscopy, which was the main focus of the previous sections, provides a way to record neuronal activity at the cellular and sub-cellular levels in a given FOV, typically limited to a few 100 μm. These recordings allow one to thoroughly investigate local microcircuits comprised from 100s to 1000s of cells at a time. At the other end of the spectrum, cortex-wide (i.e., widefield) imaging trades cellular resolution for increased FOV (millimeters) and enables exploration of the overall activity through imaging of the entire cortical surface through one-photon illumination[143] [Figs. 4(a) and 4(b)]. The captured signal exhibits significantly different statistics from micron-resolution imaging both in time and space: the spatial resolutions are too coarse to isolate activity signals of individual neurons, but instead can capture brain-wide activity patterns.[48] As a developing modality it presents many open questions regarding processing and analysis. Current approaches are nascent and inconsistent across labs. Moreover, validation and comparisons between distinct approaches are limited. In this section, we review the arising challenges and leading technological and computational approaches for capturing and processing of widefield calcium imaging.

The collection of widefield signals is performed using a scientific CMOS camera, capable of imaging hundreds of frames per second. To control the acquired data size and increase framerate, most researchers reduce the spatial resolution to 512×512 pixels, resulting in 10^2 to 10^4 μm spatial resolution.

The acquired time traces reflect an aggregated summary of the neuronal activity captured from thousands of cells, cellular compartments, and depths (although mostly superficial layers[12]). Estimation of spike rates is usually not performed for widefield signals as the captured signals may originate from various cell parts such as axons, dendrites as well as somas, each related to a different kernels. While this issue can be resolved by calcium indicators targeting specific parts of the cell at the expense of limited temporal resolution,[145,146] most researchers find the standard GCaMP indicators as the (current) best in terms of spatial and temporal resolution for capturing synaptic activity, providing a valuable tool for exploration of the dynamics of large-scale networks and their relation to complex behavior, perception, and cognition.[147,148]

Preprocessing of widefield recordings typically includes four stages: alignment, normalization, hemodynamics correction, and parcellation. Below we describe the motivation for each

Fig. 4 Widefield imaging. (a) Imaging frames of widefield signal. (b) Left—a single frame, red "*x*" correspond to motor, retrosplenial (RS), somatosensory (SS), and visual cortices, right—time-traces of the corresponding locations. (c) Left—anatomical atlas CCFv3, right—functional parcellation by LSSC. (d) Spatial components derived by SVD. Data from Ref. 144.

stage and common practices. Imaging of the entire cortical surface naturally enables analysis/ modeling across animals. To facilitate a one-to-one correspondence of data acquired from different animals, frames of each session are registered to align to a global template of the cortex, according to several anatomical control points using an affine transform.[149] In many cases, the captured time traces exhibit a slow decrease in baseline activity, ascribed to bleaching. This effect is easily removed by subtracting the slow trend (evaluated by low-pass filtering) from each pixel.[150] To equalize spatial differences of expression levels, each pixel is normalized with respect to its own overall variance post detrending.

The next stage of preprocessing aims to correct the hemodynamics artifacts, which is unique to widefield signals and is not typically present in multiphoton imaging. Fluctuations in blood flow and oxygenation alter excitation and emission of photons due to hemoglobin absorption. This phenomenon contaminates the captured signals with unwanted dynamic components.[12] The most common approach to correct this artifact is to alternate the emitted light with an additional reference channel, for example, UV light ($\sim$400 nm). As GCaMP6 is isosbestic to UV light, the emitted photons are assumed to be independent of neuronal activity whereas the data channel (typically blue $\sim$488 nm) will cause emission of photons affected by fluctuations of both neuronal activity and hemodynamics signals.

The Beer–Lambert law is an exponential model for the measured light intensity as a function of wavelength, absorption, and traveled path. Assuming that temporal deviations from the average signal are small (as they often are for widefield calcium imaging) this relation is simplified by taking a first-order estimation of the signal in a given pixel i

$$y_t(i) = \frac{F_t^{\text{blue}}(i)}{F_t^{\text{UV}}(i)} \frac{\overline{F}^{\text{blue}}(i)}{\overline{F}^{\text{blue}}(i)}, \tag{8}$$

where $\boldsymbol{F}^{\text{blue}}$ and $\boldsymbol{F}^{\text{UV}}$ are N-dimensional vectors of the recorded signals at time t through the blue and UV channels respectively and N is the number of pixels in a frame. The corrected signal, $\boldsymbol{y}_t$, is evaluated using a pixelwise linear regression.[150,151] Recently, a computational approach for improving hemodynamics reduction based on a single reference wavelength was proposed.[152] This approach exploits spatial dependencies between pixels by formulating a multivariate model

$$\boldsymbol{F}_t^{\text{blue}} = \boldsymbol{z}_t + \boldsymbol{y}_t + \nu_t \quad \boldsymbol{F}_t^{\text{UV}} = \boldsymbol{A}\boldsymbol{z}_t + \xi_t, \tag{9}$$

where $\boldsymbol{F}^{\text{blue}}, \boldsymbol{F}^{\text{UV}}, \boldsymbol{y}, \boldsymbol{z}, \nu,$ and ξ are the n-dimensional vectors, and n is the number of pixels included in a certain local patch. The corrected signal is estimated using the optimal linear predictor, $\hat{\boldsymbol{y}} = \boldsymbol{H}_1\boldsymbol{F}^{\text{blue}} + \boldsymbol{H}_2\boldsymbol{F}^{\text{UV}}$, where the matrices $\boldsymbol{H}_1, \boldsymbol{H}_2$ are evaluated from the signals at both wavelengths. Taking the patch size to be 1, this approach reduces to pixelwise regression Eq. (8) where using $n > 1$ leads to improved reduction of the hemodynamics artifact.[152] An alternative approach relies on using two reference wavelength (thus alternating three channels altogether) to obtain a more accurate correction of hemodynamic absorption.[153]

The signals at this stage are time-traces of the activity at individual pixels. The data are high-dimensional in space (typically over 10^4 pixels) and subject to several noise sources (e.g., electronic, photonic shot noise). The final stage of widefield preprocessing is therefore to extract a compact representation of brain activity and filter out the noise component. As with spatiotemporal decomposition models in TPM, most methods for extracting this representation can be formulated as a linear decomposition model: $\boldsymbol{Y} \approx \boldsymbol{A}\Phi^T$, where $\boldsymbol{Y}$ is a $N \times T$ matrix of the activity of N pixels at T time frames, $\boldsymbol{A}$ is a $N \times d$ matrix of spatial components, Φ^T is a $d \times T$ matrix of temporal components, and $d < N$. Different methods for decomposing $\boldsymbol{Y}$ vary from solely relying on anatomical features to being completely data-driven, which affects the spatial interpretability of $\boldsymbol{A}$ accordingly. Choosing one method over another should be done considering what downstream analyses will be used and the overall biological hypothesis of the research.

Using singular value decomposition (SVD) to reduce the dimension of widefield data relies on the assumption that the variance of neuronal activity within the widefield signal is significantly higher than the noise variance. The activity is decomposed to $\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ where $\boldsymbol{S}$ is a diagonal matrix of the singular values and $\boldsymbol{U}, \boldsymbol{V}$ are orthogonal matrices and a low-dimensional representation $\boldsymbol{A}\Phi^T$ is obtained by setting

$$\Phi^T = \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_d \end{pmatrix} (\boldsymbol{v}_1 \quad \ldots \quad \boldsymbol{v}_d)^T, \qquad \boldsymbol{A} = (\boldsymbol{u}_1 \quad \ldots \quad \boldsymbol{u}_d), \tag{10}$$

where $\mathbf{v}_i$ and $\boldsymbol{u}_i$ are the columns of $\boldsymbol{V}$ and $\boldsymbol{U}$, respectively and s_i is the singular values. The number of components, d, is selected to capture at least 80% to 90% of variance explained of the data, assuming that the remaining 10% to 20% relates to noise. The spatial components $\boldsymbol{u}_i$ are not constrained to be localized or non-negative and therefore the temporal components extracted in Eq. (10) do not indicate a trace of activity related to a specific brain region. Postprocessing is performed in the reduced dimension domain and then projected back to the full dimension using $\boldsymbol{U}$. For example, in Ref. 149, SVD components were used to measure how well external variables (behavior and stimuli) can predict brain activity. The trained regression parameters, each computed per temporal component, were projected to the brain-mask domain to biologically interpret the statistical findings.

A different approach is to take advantage of the spatial structure of widefield signals where adjacent pixels are typically highly correlated while noise is uncorrelated. Therefore a compact and (spatially) filtered representation can be obtained by dividing the brain into localized subregions (parcels) and extracting the average trace within each parcel. In this case $\boldsymbol{A}$ is non-negative where each row relates to a specific brain parcel with the corresponding column in Φ as its activity time trace. Unlike imaging of local circuits, where identifying cell boundaries is a well-defined task (although not simple for automation), detection of parcel boundaries is not straightforward.

The most common approach for cortex-wide parcellation is to use a predefined atlas based on anatomical features such as the common coordinate framework (CCFv3) proposed by the Allen Institute for Brain Research[154] or the mouse brain atlas of Paxinos and Franklin[155] [Fig. 4(c)]. Parcellation based on anatomical atlases presents many advantages—each brain parcel represents a well-known biological functionality (e.g., vision, motor, and sensory) and a straightforward way to compare neuronal activity across animals (and studies). However, it is often observed that the spatial patterns of activity in some regions are not well described by anatomical outlines [e.g., via principal components analysis Fig. 4(d)]. Identification of subcortical borders can be performed experimentally by presenting sensory stimuli, e.g., the auditory cortex[156] or the visual cortex.[157] These methods are highly efficient for detecting boundaries of subregions within a specific cortical area (visual cortex and auditory cortex), but cannot be used to detect regions that are not responsive to sensory stimuli. Therefore, computational methods for functional parcellation [Fig. 4(c)] have been a major target of research in recent years for calcium imaging as well as in the fMRI community.[158]

Localized seminonnegative matrix factorization (LocaNMF) is a recently proposed approach aiming to tackle this issue by formulating an optimization problem for minimizing the mean-squared error between the widefield signal $\boldsymbol{Y}$ and the estimated signal $\boldsymbol{A}\Phi^T$ such that the columns of $\boldsymbol{A}$ are non-negative and localized according to anatomical clusters.[34] LocaNMF produces, by nature, spatial patterns that are similar to the anatomical boundaries used by the optimization process and therefore typically does not deviate much from the anatomical atlas. Related to NMF, a linear one-hidden-layer autoencoder has also been used to identify parcellations in auditory cortex.[159] Similar in spirit, GraFT,[35] being agnostic to spatial morphology due to its underlying graph-based modeling, has also been applied to extracting (potentially overlapping) widefield spatial maps in rat and mouse macroscopic data. In a recent study, functional parcellation took a different turn by adding a temporal component to brain parcellation. This approach is based on finding repeated spatiotemporal patterns of activation termed motifs, which can be viewed as a time-varying brain parcel. The overall activity is therefore represented as a sum of convolution terms between each motif and its corresponding time trace.[160]

A different approach for brain parcellation is to cluster the brain into regions of coactivity, with no regard to anatomical features. In this case, the matrix $\boldsymbol{A}$ is comprised of binary vectors, each corresponding to a specific brain parcel. Li et al.[161] proposed an iterative greedy algorithm for parcellating the brain based on correlation similarity. Other approaches use correlations as graph weights connecting pixels serving as nodes. Parcels are obtained by clustering the graph

using Ncut, where the number of parcels is a hyperparameter,[162] or a greedy adaptation of spectral clustering, where the number of parcels is learned from the data.[163]

Overall, the product of functional parcellation methods describe the spatial distribution of coactivity in a given session (per animal). These patterns may be consistent across sessions[163] but are not, in general, uniform across animal and obviously vary with the nature of the experiment (e.g., spontaneous activity and task directed). In that regard, as performed for the SVD approach, postprocessing values produced per functional parcel (e.g., goodness of fit and modeling coefficients) can be projected onto the brain mask using the spatial components.

Mapping of cortical neuronal activity is also recently addressed experimentally through multimodal imaging where widefield calcium imaging recording is performed simultaneously with fRMI signals[164] or imaging of local circuit using two-photon microscopy[150] or electrophysiology[165–168] where specific regions are identified as highly correlated to a specific cell (or subpopulation).

6 Modern Challenges in Optical Functional Imaging

While there remain numerous challenges in widespread multiphoton and widefield imaging, there are further emerging challenges that will necessitate new data processing solutions in the near future. We outline here a number of primary avenues that we believe show promise but have many outstanding problems to solve.

6.1 Imaging Morphologies Beyond the Soma

More recently, variants of optical imaging have aimed to expand the scope of accessible brain signals by imaging both larger and smaller neural structures. At one end of this spectrum, zooming in enables the imaging of dendritic and spine structures, which captures how individual neurons communicate.[169–174] Dendritic[175] and axonal[176] imaging, while also having sparse temporal statistics, can have long, thin spatial profiles that span the entire FOV. We note that these types of neural morphologies are also vital in some species, such as *Drosophila melanogaster*, where dendritic activity is vital to tracking neural processing.[177,178] Approaches to dendritic/axonal imaging have largely followed the path of somatic imaging analysis. For example, recent versions of Suite2p[69] can be run in "dendrite mode." The long, stringy morphologies, however, are at odds with typical built-in assumptions of spatial locality. A more recent approach instead redefines pixel ordering on a data-driven graph to better identify irregular morphologies.[35] Astrocytes also exhibit irregular morphologies and size, for which targeted approaches have been recently developed.[179,180] Interestingly, larger-scale optical imaging of hemodynamic activity also captures components (i.e., blood vessels) that have spatial statistics more similar to dendrites than somas.[181]

6.2 Voltage Imaging

Voltage imaging is a technique that has been used for decades to record changes in neural activity using voltage-sensitive dyes.[182] Voltage imaging, as compared with recording changes in calcium, is a more direct way of measuring neural signals. A comparison of population responses to optical recordings using calcium indicators and voltage indicators showed major differences in the temporal response of the recorded calcium signal as compared with the voltage signal.[183] Widefield voltage imaging has also been explored[184] and faces many of the same challenges as widefield imaging with calcium sensors (see Sec. 4.6).

Technology for the optical recording of voltage signals has improved rapidly over the past few years. The development of improved voltage sensors in the form of bright genetically encoded voltage indicators (GEVIs) has enabled high-resolution voltage recordings at multiple spatial scales.[185] Genetic targeting of these indicators to subcellular structures isolate signal to particular neuronal structures, further increasing the signal-background ratio (SBR). These improvements have allowed researchers to generate optical voltage recordings from a population of cells in awake, behaving animals.[186]

Several challenges still exist that prevent the generation of large-scale optical voltage recordings. Voltage indicators are membrane-bound, which limits the total concentration of the sensors is compared with calcium indicators, which may fill the whole cytoplasm. Generally, GEVIs are, at least for now, dimmer than their GECIs counterparts. To take advantage of the improved temporal response function, recordings must be made at much higher framerates as compared with calcium indicators ($\sim$1000versus $\sim$30 Hz). The combination of these challenges reduces the overall spatial scale of current cellular or subcellular resolution recordings with voltage indicators.

Current-voltage imaging analysis includes both matrix factorization[91,187] and deep learning[52] demixing approaches similar to methods used in calcium imaging. The potential for low-SNR and non-Gaussian noise statistics can complicate demixing. Moreover, extremely high temporal resolutions create much larger datasets, increasing the computational cost of processing. Finally, non-negativity is a basic assumption built into many calcium imaging analysis methods: i.e., deviations from baseline are only positive, however, voltage traces have no such constraint.

6.3 *Computational Imaging*

One critical avenue that may completely upend much of how optical imaging data is processed is computational imaging. Computational imaging represents a paradigm wherein optical and algorithmic components are codesigned to compensate and enhance each other and achieve superior results to advances in either area independently.[188,189] Codesigned approaches are nascent in *in-vivo* functional imaging of the brain, with only a few examples aimed at faster imaging[41,190] or volumetric imaging.[16,45,191,192] The algorithmic designs for computational imaging tend to require specialized and often unique processing elements that invert the optical path of the co-designed microscope.[189] Examples include tomography,[41] light-field imaging,[45,192–194] combined light-field and tomography,[191] stereoscopy,[16] and computational systems for imaging through scattering tissue.[195]

While basic denoising or other discussed techniques might still be applicable, more highly coded optics may not even be able to use the same motion correction, let alone the other advances in demixing of individual neuronal signals. For example, lightfield and stereoscopy methods need demixing methods for decoding projections of volumetric components onto 2D arrays. Thus, new frontiers are constantly expanding and requiring novel advances in our handling of functional optical data.

7 Validation and Assessment

One of the most difficult tasks in creating widely applicable and robust calcium image processing methods is proper assessment.[196] The effects of mismatch between the necessary simplified statistical assumptions in the signal processing models and the actual data properties must be explored in terms of the effect on the fidelity of extracted signals and, when possible, the later scientific analyses. A prime example of such an effect is explored in Ref. 92, in which it is shown that the *i.i.d.* Gaussian noise assumption often considered can create bleed-through between overlapping cells and additional fluorescent biological processes in the tissue. The result is that the time traces have high levels (between 15% and 25%) of transient events being false transients in that they do not reflect the activity of that cell, but rather other, nearby, fluorescing components in the tissue. While significant in and of itself, it is further noted that these errors can cause errors in the interpretation of the neural activity, including the skewed discovery of location encoding in the hippocampus.[92]

To identify the accuracy of optical imaging processing algorithms, a number of avenues have emerged. Specifically, four current available avenues are assessment based on (1) manual annotation, (2) biophysical simulation, (3) local self consistency of global decompositions, and (4) consistency based on external experimental variables.

We note that a fifth form of validating signals extracted from optical measurements take the form of simultaneous electrophysiological recording and optical imaging.[130,137,197] The number of simultaneous cells that can be recorded during imaging within an FOV, however, is limited. Thus these recordings have primarily taken a role in assessing either the accuracy of estimated

spikes from optically recorded calcium,[130,197] or the similarity between electrically and optically recorded neural signals.[137] Until now, this approach does not meet the scale required to completely assess the processing of a full FOV.

7.1 *Manual Annotation*

The most basic assessments of optical imaging analysis are the manual labeling of cells. This is typically done by annotating an image summarizing the structure in an imaging session, i.e., temporal max projection, temporal average, local temporal correlation,[198] or nonlinear embeddings.[199] We note that while manual labeling is typically done on the processed data, labels can also be obtained via anatomical imaging (e.g., *z*-stacks or nuclear labeling with activity-invariant fluorescent proteins). Comparison to manual annotation gives clear metrics: hits (manually identified cells that overlapped significantly with a match in the returned spatial profiles) and misses (those with no match) [Fig. 5(a)].

However, there are limitations to assessment via manual annotation. Time traces are not available as ground truth and must be inferred from the data given in the annotations. Additionally, the annotations are often incomplete, excluding sparsely firing cells, or nonsomatic components. The effect is that found profiles that do not match well to the manual annotations could appear as "false alarms" but fall into many categories. They may be actual cells in the data missed by the annotator, they may be algorithmic artifacts caused by merging or splitting parts of cells,[92] or they may be overfitting to noise or neuropil. Thus at best, manual annotations give a lower bound on true hits and missed detections but not much information on false positives. Some of the limitations stemming from human error can be removed in new datasets consisting of electron microscopy reconstructions of the imaged tissue.[51] This dataset, and any that follow,
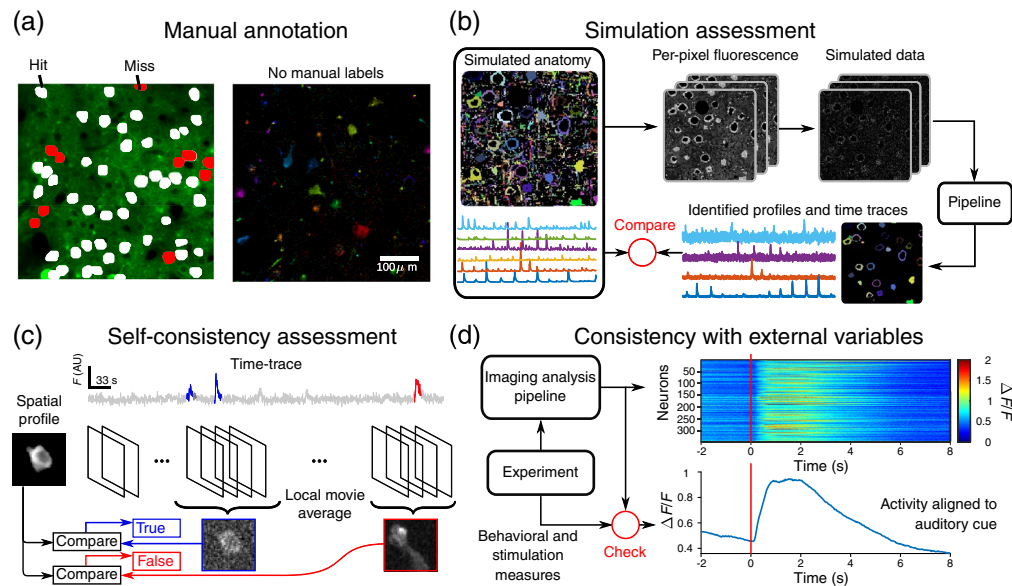
Fig. 5 Assessing optical imaging analysis outputs. (a) Manual annotation, e.g., the NeuroFinder benchmark dataset (example shown here), can help detect true detections (white) and missed cells (red), as compared with a labeling of the spatial area imaged (right). Manual annotation, however, does not provide information on cells that do not match any of the manual annotations (left). (b) Simulation-based assessment uses *in silico* simulations of activity and anatomy to create faux fluorescence microscopy data. The outputs of an analysis pipeline using the data can be compared with the simulated ground truth. (c) Self-consistency measures use the global demixed components and analyze local space-time extents of the movie to check if the spatial and temporal components still describe the data well. For example, is activity erroneously attributed to a found cell. (d) Consistency with external variables examines if expected gross behavior of the imaged neural populations match behavioral or stimulation events corecorded with the fluorescence video. For example, shown here are neurons in motor cortex responding to a motion made by a mouse responding to an auditory cue.[200]

provide anatomical ground truth in the form of a registerable volume to match spatial components to, although time traces are still unavailable.

7.2 *Biophysical Simulations*

Another important form of assessment is *in-silico* simulation.[123,201] Simulation plays an important role in assessing the fundamental limits of algorithms across signal processing and machine learning. In particular, simulations are vital when ground truth information is difficult to obtain, either efficiently, or at all. Functional fluorescence microscopy is exactly one such situation. The time, effort, and expense of simultaneous recording fluorescence imaging and electrophysiology is a remarkable effort for only a limited portion of the ground-truth data. Simulations offer a potential solution by leveraging anatomical and physical knowledge of the system to generate data where the underlying activity and anatomy driving the synthetic observations are completely known and can be compared with Fig. 5(b).

Simulations, however, pose their own risk. Simplistic simulations can miss complexities in the real-data imaging statistics, e.g., non-Gaussian or non-*i.i.d.* noise. Complex simulations run the opposite risk of being so detailed that the computational run-time and memory requirements become excessive and capture details irrelevant to the assessment being conducted. Recent work on the neural anatomy and optical microscopy (NAOMi) simulator seeks to balance these competing needs.[123] To date, NAOMi has been applied to a number of scenarios, such as testing different demixing algorithms,[35] denoising methods,[32] sensitivity to negative transients,[202] and testing/training calcium imaging demixing algorithms.[131]

7.3 *Data Consistency*

A third form of assessment uses no ground truth data, manual or synthetic, and instead focuses on self-consistency of the data model. The spatial profiles and time traces give a global decomposition of the data by minimizing data fidelity and regularization terms over the entire dataset. In this approach, one can focus on a smaller segment of the data and check if the global decomposition matches the local statistics. Specifically, one can check if in the movie frames, during which a given cell was purported to have fired (a transient burst in the time trace during those frames), if the shape of the cell truly appears in the video. Recent work used this local averaging idea to identify how many algorithms are not locally consistent: i.e., demonstrating that activity from different sources bleed into each other[92] [Fig. 5(c)]. The resulting errors (termed false transients) can influence scientific findings and can appear in the time-trace estimates of many different algorithms and across many datasets.[92] To address this finding, additional work has sought to develop more robust time-trace estimators that prevent these errors.[87,92,94]

7.4 *Consistency with External Measures*

The final form of validation we discuss is with respect to external measures outside the imaged brain area. To explore the relation between brain activity and behavior, perception, and cognition most experimental setups record, simultaneously with the brain activity, external variables such as spontaneous behavior (whisking, running, and pupil size), responses to sensory stimuli, and task-related behavior. A thorough examination of the extracted traces of activity with respect to external variables can be an important tool for the assessment of fluorescence microscopy processing. For example, aligning the activity traces to external onsets such as repeated presentations of a sensory cue, specifically trained behavior or even running onsets allows one to examine brain activity in a behavioral context. Applying this strategy should be done with caution, as uninstructed behavior may cause significant cell-to-cell and trial-to-trial variability.[149] Still, averaging across dozens of well-pronounced presentations (visual or auditory cues for example) and across cells usually yields a significant increase of activity in the visual/auditory cortices[203,204] or to a distinct pattern of activation of the motor cortex in well-trained animals[200] [Fig. 5(d)]. For widefield imaging, these same strategies can be applied with respect to the appropriate brain parcels. If no sensory cues are presented, averaging across running onsets should lead to a significant increase in the overall activity of the cortex.[152] Alternatively, additional modalities of

neuronal activity can be used for validation of calcium imaging signals such as electrophysiology,[115] fMRI,[164] or dual imaging of widefield and two-photon imaging.[150]

8 Discussion

We have aimed to review here a large portion of the literature related to the analysis of functional optical microscopy data. The topics we have covered aim to provide a practical overview of how optics choices affect signal processing challenges, and the many methods that have been developed to solve these challenges. While broad categories, such as denoising, motion correction, and in particular signal extraction, have been explored in detail in the literature, there are many other challenges that have yet to be solved.

For one, robust alignment across sessions is an ongoing challenge. While matching cells based on anatomical morphology is currently possible, effects such as nonlinear shearing in the brain and axial drift can create situations where only portions of an FOV can be recovered and aligned across recording sessions. Even within long sessions, these effects can remove cells from parts of the recording or reduce the signal quality. Identifying the periods where cells are not visible is critical in determining how subsequent analyses should interpret zero values. Passing zero values for these times into typical analyses assumes a cell not firing, rather than the neuron's activity being unavailable. Instead, missing data methods must be employed, which need to know exactly when the missing data occurred. As chronic recordings become more commonplace, we expect that cross-session alignment and axial shift compensation will become critical hurdles to pass.

Another growing area of interest is the real-time analysis of functional microscopy data. Real-time analysis enables closing of the loop, i.e., the ability to use estimates of neural activity to drive future experimental trials.[205] While basic manual annotation is trivial to move to an online setting, full demixing algorithms that solve many of the aforementioned challenges are still in their nascent stages. Initial work is promising in the ability to motion correct[104] and infer calcium activity.[206,207] However, the ability to infer activity and cell shapes completely online for dense neural fields is still an ongoing research direction.

A critical aspect of analysis methods that we have not discussed here is the computational cost across methods. There are large variations in cost from simple averaging to training entire deep neural networks. Unfortunately in some regards, these disparities match the high variability in the computational infrastructure available across labs. Moreover, as optical imaging of neurons continues to advance, computationally efficient techniques will only become more critical. Already at moderate fields of view, there are up to 10^3 neurons. New methods leveraging the latest microscope designs are recording $10^5 - 10^6$ neurons at once.[10,11] Furthermore, volumetric (3D) imaging complicates image analysis by rendering many planar segmentation methods inapplicable.[16,38,208-213] In these cases new methods will need to be developed, and computational efficiency will be key to analyzing these very high dimensional datasets.

Another topic we have not elaborated upon is the effect of indicators on signal analysis. We have focused primarily on generic properties of fast- and slow- calcium indicators and, as a faster comparison, voltage indicators. Most indicators share basic characteristics with these classes. In the spirit of universal analysis pipelines, one possible approach is to further abstract all indicators into basic quantities, e.g., quantum efficiency, coupling affinity, etc., which can be used to tune parameters in more generic versions of the current algorithms. This approach, which requires a precise estimate of these quantities and their variation, would greatly benefit the user base.

Another change with different indicators is the assumption of non-negativity. Calcium imaging analysis has largely assumed only positive deviations from baseline. Voltage imaging and widefield calcium imaging all display negative dips as well, which will represent more degrees of changes that need to be implemented into more general pipelines. We note that even for calcium imaging, interestingly new explorations discuss the presence of so-called negative transients, as well as the ability of various algorithms to cope with these unexpected dips.[202]

One increasingly popular group of approaches in the analysis of functional optical imaging is to leverage advances in deep learning. Critical to the success of deep learning approaches is (1) the availability of training data and (2) the generalization of the trained system to new

datasets. For training data, most systems use NeuroFinder[197] and/or the Allen Institute Mouse Brain Observatory.[214] However, both datasets suffer from both mislabeled data and missing labels, or erroneous time-trace estimates.[92] Generalization is tougher to ensure, as image statistics can affect deep learning systems in a myriad of unexpected ways. Therefore extensive experimentation is required, e.g., by testing a trained system on imaging from different depths to explore the effect of tissue distortion.[84] Thus, despite the fact that pre-trained networks are fast to run, these challenges create a high level of uncertainty in using pretrained networks accurately. Individual labs can instead choose to annotate their data to train networks from scratch, such that they are optimized for the imaging used locally. While this solves some of the aforementioned challenges (assuming care is taken in annotation), the training procedure can be computationally intensive relative to a lab's computational capabilities.

One path to improving training data has been to augment the dataset, such as using random rotations and reflections.[215] For optical imaging, augmentation should use the degrees of variation known through characterizations of optical-tissue interactions. The use of biophysical simulators to generate synthetic training data, or even modulate real data, can prove useful. In fact, recent work used the NOAMi simulation suite[123] to generate data for training a spike estimation network.[131] Finally, another consideration in deep learning approaches is a heavy class imbalance both spatially and temporally between neurons and background in functional fluorescence imaging datasets, i.e., background frequently dominates the acquired FOV, and the desired neural activity may be temporally sparse.

As we have noted, assessment of data segmentation quality is, in general, challenging. Evaluation datasets must reflect typical use cases of the community, and therefore they need to constantly be updated and expanded. Due to the rapid expansion of functional microscopy technology, the typical use case is already eclipsing standard datasets, requiring the description of new benchmarks. New datasets are beginning to fill the void, e.g., in voltage imaging.[52] However, the quality and diversity of imaging data are only accelerating. For example, it would be immensely valuable to the community to provide benchmark data for especially difficult situations, such as cortex-wide widefield data. Even more powerful would be an amalgamation of different benchmark datasets—both real and synthetic—similar to what SpikeForest[216] has done for spike sorting algorithms. Only with such a wide-ranging set of tests will the strengths and weaknesses of different algorithms become apparent.

Assessment is even more difficult in widefield data, where anatomy provides minimal aid in assessing the validity of spatial component shapes. Neighboring brain areas can share tracts of coordinated activity, and activity can likewise be constrained to a small portion of a brain area. Parcellations thus need to be carefully considered in the context of the behavior and other recordings. In this line of consideration is the nature of parcellations themselves. An unsolved question currently facing the community is whether parcels can overlap: current methods include both solutions that allow[34,35,159] or prohibit[161–163] overlapping parcels. Both confer different interpretations: nonoverlapping parcels provide a solution wherein each component represents a given brain area's activity while overlapping can enable the time traces to be better event- or behavior-locked in the case of a brain region having multiple uses. Future work should explore the interplay between these two windows into widefield data.

One final major concern is the reproducibility and accessibility of functional imaging analysis techniques (which we note is not restricted to this modality[217]). Making code available and easier to use[33,218] is only the first step in this path. Learning the intricacies in robustly running the software in new hardware environments, such as local clusters or desktops, is a hard-earned expertise. One option is to ensure that the software is robust and tested on many systems. This approach requires either systems engineering skills that are not typically within the budget or scope of a typical research lab. Instead, a concerted effort across an entire community is required. The community has to coalesce around—and contribute actively to—a specific approach. An alternative being explored is to have individual labs containerize their software, i.e., create shareable virtual environments that are tested with the given software. In neuroscience, an emerging example is the NeuroCAAS system,[219] which provides Dockerized implementations of algorithms to be run, e.g., on the amazon web service.

In conclusion, functional optical imaging in neuroscience is rapidly growing and is accurate, and automated processing of the massive data being generated is becoming increasingly essential

to the continued progress of understanding the brain. Solving these challenges will take both forms of methods that enable, e.g., real-time, robust, fast analyses, and well-engineered infrastructure that democratizes the current advances to the growing number of labs employing functional microscopy in their experiments. We thus expect that this area will continue to grow rapidly in the next decade, drawing on increased interest from labs across neuroscience, data science, imaging, and other related disciplines.

Disclosures

The authors declare no conflicts of interest in the preparation and publication of this work.

Acknowledgment

References

1. J. J. Jun et al., "Fully integrated silicon probes for high-density recording of neural activity," *Nature* **551**(7679), 232 (2017).
2. N. A. Steinmetz et al., "Neuropixels 2.0: a miniaturized high-density probe for stable, long-term brain recordings," *Science* **372**(6539), eabf4588 (2021).
3. E. Mace et al., "Functional ultrasound imaging of the brain: theory and basic principles," *IEEE Trans. Ultrason. Freq. Control* **60**(3), 492–506 (2013).
4. A. Urban et al., "Real-time imaging of brain activity in freely moving rats using functional ultrasound," *Nat. Methods* **12**(9), 873–878 (2015).
5. D. Y. Takahashi et al., "Social-vocal brain networks in a non-human primate," bioRxiv (2021).
6. A. E. Hady et al., "Chronic brain functional ultrasound imaging in freely moving rodents performing cognitive tasks," bioRxiv (2022).
7. M. Tik et al., "Ultra-high-field fMRI insights on insight: neural correlates of the Aha!-moment," *Hum. Brain Mapp.* **39**(8), 3241–3252 (2018).
8. O. Viessmann and J. R. Polimeni. "High-resolution fMRI at 7 Tesla: challenges, promises and recent developments for individual-focused fMRI studies," *Curr. Opin. Behav. Sci.* **40**, 96–104 (2021).
9. L. Huber et al., "Ultra-high resolution blood volume fMRI and BOLD fMRI in humans at 9.4 T: capabilities and challenges," *Neuroimage* **178**, 769–779 (2018).
10. M. B. Bouchard et al., "Swept confocally-aligned planar excitation (SCAPE) microscopy for high-speed volumetric imaging of behaving organisms," *Nat. Photonics* **9**(2), 113–119 (2015).
11. J. Demas et al., "High-speed, cortex-wide volumetric recording of neuroactivity at cellular resolution using light beads microscopy," *Nat. Methods* **18**(9), 1103–1111 (2021).
12. Y. Ma et al., "Wide-field optical mapping of neural activity and brain hemodynamics: considerations and novel approaches," *Philos. Trans. R. Soc. B Biol. Sci.* **371**(1705), 20150360 (2016).
13. A. O. Nunez-Elizalde et al., "Neural basis of functional ultrasound signals," bioRxiv (2021).
14. T. Xie et al., "GRIN lens rod based probe for endoscopic spectral domain optical coherence tomography with fast dynamic focus tracking," *Opt. Express* **14**(8), 3238–3246 (2006).
15. C. Wang and N. Ji, "Characterization and improvement of three-dimensional imaging performance of GRIN-lens-based two-photon fluorescence endomicroscopes with adaptive optics," *Opt. Express* **21**(22), 27142–27154 (2013).
16. A. Song et al., "Volumetric two-photon imaging of neurons using stereoscopy (vTwINS)," *Nat. Methods* **14**(4), 420 (2017).

17. D. R. Beaulieu et al., "Simultaneous multiplane imaging with reverberation multiphoton microscopy," arXiv:1812.05162 (2018).
18. W. Shao et al., "Wide field-of-view volumetric imaging by a mesoscopic scanning oblique plane microscopy with switchable objective lenses," *Quant. Imaging Med. Surg.* **11**(3), 983–997 (2020).
19. A. Badura et al., "Fast calcium sensor proteins for monitoring neural activity," *Neurophotonics* **1**(2), 025008 (2014).
20. M. Kannan et al., "Fast, in vivo voltage imaging using a red fluorescent indicator," *Nat. Methods* **15**(12), 1108 (2018).
21. Y. Shen et al., "Genetically encoded fluorescent indicators for imaging intracellular potassium ion concentration," *Commun. Biol.* **2**(1), 18 (2019).
22. T.-W. Chen et al., "Ultrasensitive fluorescent proteins for imaging neuronal activity," *Nature* **499**(7458), 295 (2013).
23. H. Dana et al., "High-performance calcium sensors for imaging activity in neuronal populations and microcompartments," *Nat. Methods* **16**(7), 649–657 (2019).
24. W. Denk, J. H. Strickler, and W. W. Webb, "Two-photon laser scanning fluorescence microscopy," *Science* **248**(4951), 73–76 (1990).
25. J. C Jung et al., "In vivo mammalian brain imaging using one-and two-photon fluorescence microendoscopy," *J. Neurophysiol.* **92**(5), 3121–3133 (2004).
26. C. Rodriguez et al., "Three-photon fluorescence microscopy with an axially elongated bessel focus," *Opt. Lett.* **43**(8), 1914–1917 (2018).
27. N. J. Sofroniew et al., "A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging," *Elife* **5**, e14472 (2016).
28. R. Yasuda et al., "Supersensitive RAS activation in dendrites and spines revealed by two-photon fluorescence lifetime imaging," *Nat. Neurosci.* **9**(2), 283 (2006).
29. R. Homma et al., "Wide-field and two-photon imaging of brain activity with voltage and calcium-sensitive dyes," in *Dynamic Brain Imaging*, F. Hyder, ed., pp. 43–79, Springer (2009).
30. A. D. Jacob et al., "A compact head-mounted endoscope for in vivo calcium imaging in freely behaving mice," *Curr. Protoc. Neurosci.* **84**(1), e51 (2018).
31. L. Zhang et al., "Miniscope GRIN lens system for calcium imaging of neuronal activity from deep brain structures in behaving animals," *Curr. Protoc. Neurosci.* **86**(1), e56 (2019).
32. J. Lecoq et al., "Removing independent noise in systems neuroscience data using deepinterpolation," *Nat. Methods* **18**(11),1401–1408 (2021).
33. D. A. Cantu et al., "Ezcalcium: open-source toolbox for analysis of calcium imaging data," *Front. Neural Circuits* **14**, 25 (2020).
34. S. Saxena et al., "Localized semi-nonnegative matrix factorization (locanmf) of widefield calcium imaging data," *PLoS Comput. Biol.* **16**(4), e1007791 (2020).
35. A. S. Charles et al., "Graft: graph filtered temporal dictionary learning for functional neural imaging," *IEEE Trans. Image Process.* **31**, 3509–3524 (2022).
36. C. Stosiek et al., "In vivo two-photon calcium imaging of neuronal networks," *Proc. Natl. Acad. Sci. U. S. A.* **100**(12), 7319–7324 (2003).
37. W. R. Zipfel, R. M. Williams, and W.W. Webb, "Nonlinear magic: multiphoton microscopy in the biosciences," *Nat. Biotechnol.* **21**(11), 1369–1377 (2003).
38. R. Lu et al., "Video-rate volumetric functional imaging of the brain at synaptic resolution," *Nat. Neurosci.* **20**(4), 620–628 (2017).
39. D. G. Ouzounov et al., "In vivo three-photon imaging of activity of GCaMP6-labeled neurons deep in intact mouse brain," *Nat. Methods* **14**(4), 388–390 (2017).
40. G. Katona et al., "Fast two-photon in vivo imaging with three-dimensional random-access scanning in large tissue volumes," *Nat. Methods* **9**(2), 201–208 (2012).
41. A. Kazemipour et al., "Kilohertz frame-rate two-photon tomography," *Nat. Methods* **16**(8), 778–786 (2019).
42. T. Zhang et al., "Kilohertz two-photon brain imaging in awake mice," *Nat. Methods* **16**(11), 1119–1122 (2019).
43. N. Vladimirov et al., "Light-sheet functional imaging in fictively behaving zebrafish," *Nat. Methods* **11**(9), 883–884 (2014).

44. O. Skocek et al., "High-speed volumetric imaging of neuronal activity in freely moving rodents," *Nat. Methods* **15**(6), 429–432 (2018).

45. Z. Zhang et al., "Imaging volumetric dynamics at high speed in mouse and zebrafish brain with confocal light field microscopy," *Nat. Biotechnol.* **39**(1), 74–83 (2021).

46. K. K. Ghosh et al., "Miniaturized integration of a fluorescence microscope," *Nat. Methods* **8**(10), 871–878 (2011).

47. D. J. Cai et al., "A shared neural ensemble links distinct contextual memories encoded close in time," *Nature* **534**(7605), 115–118 (2016).

48. 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).

49. A. Giovannucci et al., "CaImAn an open source tool for scalable calcium imaging data analysis," *Elife* **8**, e38173 (2019).

50. A. Aggarwal et al., "Glutamate indicators with improved activation kinetics and localization for imaging synaptic transmission," bioRxiv (2022).

51. P. Zhou et al., "Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data," *Elife* **7**, e28728 (2018).

52. C. Cai et al., "VolPy: automated and scalable analysis pipelines for voltage imaging datasets," *PLoS Comput. Biol.* **17**(4), e1008806 (2021).

53. X. Chen et al., "Brain-wide organization of neuronal activity and convergent sensorimotor transformations in larval zebrafish," *Neuron* **100**(4), 876–890.e5 (2018).

54. P Ellinger, "Fluorescence microscopy in biology," *Biol. Rev.* **15**(3), 323–347 (1940).

55. C. B. Kayasandik and D. Labate, "Improved detection of soma location and morphology in fluorescence microscopy images of neurons," *J. Neurosci. Methods* **274**, 61–70 (2016).

56. N. Korfhage et al., "Detection and segmentation of morphologically complex eukaryotic cells in fluorescence microscopy images via feature pyramid fusion," *PLOS Comput. Biol.* **16**(9), e1008179 (2020).

57. Y. K. T. Xu et al., "Automated in vivo tracking of cortical oligodendrocytes," *Front. Cell. Neurosci.* **15**, 667595 (2021).

58. L. B. Cohen and B. M. Salzberg, "Optical measurement of membrane potential," *Rev. Physiol. Biochem. Pharmacol.* **83**, 35–88 (1978).

59. A. Grinvald, "Real-time optical mapping of neuronal activity: from single growth cones to the intact mammalian brain," *Annu. Rev. Neurosci.* **8**(1), 263–305 (1985).

60. 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).

61. J. P. Cunningham and M. Y. Byron, "Dimensionality reduction for large-scale neural recordings," *Nat. Neurosci.* **17**(11), 1500–1509 (2014).

62. G. Mishne et al., "Hierarchical coupled-geometry analysis for neuronal structure and activity pattern discovery," *IEEE J. Sel. Top. Signal Process.* **10**(7), 1238–1253 (2016).

63. H. Benisty et al., "Rapid fluctuations in functional connectivity of cortical networks encode spontaneous behavior," bioRxiv (2021).

64. S. Vyas et al., "Computation through neural population dynamics," *Annu. Rev. Neurosci.* **43**, 249–275 (2020).

65. J. W. Pillow et al., "Spatio-temporal correlations and visual signalling in a complete neuronal population," *Nature* **454**(7207), 995–999 (2008).

66. 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).

67. E. A. Pnevmatikakis et al., "Simultaneous denoising, deconvolution, and demixing of calcium imaging data," *Neuron* **89**(2), 285–299 (2016).

68. B. D Haeffele and R. Vidal, "Structured low-rank matrix factorization: global optimality, algorithms, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.* **42**(6), 1468–1482 (2020).

69. M. Pachitariu et al., "Suite2p: beyond 10,000 neurons with standard two-photon microscopy," bioRxiv (2016).

70. M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE* **98**(6), 972–982 (2010).

71. B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature* **381**(6583), 607 (1996).

72. M. Pachitariu et al., "Extracting regions of interest from biological images with convolutional sparse block coding," in *Adv. Neural Inf. Process. Syst.*, pp. 1745–1753 (2013).

73. F. Diego and F. A. Hamprecht, "Sparse space-time deconvolution for calcium image analysis," in *NIPS*, pp. 64–72 (2014).

74. A. Petersen, N. Simon, and D. Witten, "Scalpel: extracting neurons from calcium imaging data," *Ann. Appl. Stat.* **12**(4), 2430 (2018).

75. A. Klibisz et al., "Fast, simple calcium imaging segmentation with fully convolutional networks," *Lect. Notes Comput. Sci.* **10553**, 285–293 (2017).

76. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).

77. P. Kaifosh et al., "Sima: Python software for analysis of dynamic fluorescence imaging data," *Front. Neuroinf.* **8**, 80 (2014).

78. Q. Spaen et al., "HNCcorr: a novel combinatorial approach for cell identification in calcium-imaging movies," *eNeuro* **6**(2), 1–19 (2019).

79. G. Mishne et al., "Automated cellular structure extraction in biological images with applications to calcium imaging data," bioRxiv 313981 (2018).

80. S. Reynolds et al., "ABLE: an activity-based level set segmentation algorithm for two-photon calcium imaging data," *eNeuro* **4**(5), 1–13 (2017).

81. T. Chan and L. Vese, "An active contour model without edges," *Lect. Notes Comput. Sci.* **1682**, 141–151 (1999).

82. S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations," *J. Comput. Phys.* **79**(1), 12–49 (1988).

83. N. Apthorpe et al., "Automatic neuron detection in calcium imaging data using convolutional networks," in *Adv. Neural Inf. Process. Syst.*, Vol. **29**, pp. 3270–3278 (2016).

84. 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).

85. Y. Bao et al., "Segmentation of neurons from fluorescence calcium recordings beyond real time," *Nat. Mach. Intell.* **3**, 590–600 (2021).

86. E. Kirschbaum, A. Bailoni, and F. A. Hamprecht, "Disco: deep learning, instance segmentation, and correlations for cell segmentation in calcium imaging," *Lect. Notes Comput. Sci.* **12265**, 151–162 (2020).

87. H. Inan, M. A. Erdogdu, and M. Schnitzer, "Robust estimation of neural signals in calcium imaging," in *NIPS*, pp. 2905–2914 (2017).

88. R. Maruyama et al., "Detecting cells using non-negative matrix factorization on calcium imaging data," *Neural Netw.* **55**, 11–19 (2014).

89. G. Mishne and A. S. Charles, "Learning spatially-correlated temporal dictionaries for calcium imaging," in *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP-2019)*, IEEE, pp. 1065–1069 (2019).

90. L. M. Tran et al., "Automated curation of cnmf-e-extracted roi spatial footprints and calcium traces using open-source automl tools," *Front. Neural Circuits* **14**, 42 (2020).

91. E. K. Buchanan et al., "Penalized matrix decomposition for denoising, compression, and improved demixing of functional imaging data," arXiv:1807.06203 (2018).

92. J. L Gauthier et al., "Detecting and correcting false transients in calcium imaging," bioRxiv 473470 (2018).

93. X.-X. Wei et al., "A zero-inflated gamma model for deconvolved calcium imaging traces," arXiv:2006.03737 (2020).

94. J. Denis et al., "Deepcinac: a deep-learning-based python toolbox for inferring calcium imaging neuronal activity based on movie visualization," *eNeuro* **7**(4), 1–15 (2020).

95. C. Stringer and M. Pachitariu, "Computational processing of neural recordings from calcium imaging data," *Curr. Opin. Neurobiol.* **55**, 22–31 (2019).

96. S. Laffray et al., "Adaptive movement compensation for in vivo imaging of fast cellular dynamics within a moving tissue," *PloS One* **6**(5), e19928 (2011).

97. 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).

98. F. C. Collman, *High Resolution Imaging in Awake Behaving Mice: Motion Correction and Virtual Reality*, Princeton University (2010).

99. K. J. Sekiguchi et al., "Imaging large-scale cellular activity in spinal cord of freely behaving mice," *Nat. Commun.* **7**(1), 1–13 (2016).

100. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Opt. Lett.* **33**(2), 156–158 (2008).

101. A. Dubbs, J. Guevara, and R. Yuste, "Moco: fast motion correction for calcium imaging," *Front. Neuroinf.* **10**, 6 (2016).

102. A. Mitani and T. Komiyama, "Real-time processing of two-photon calcium imaging data including lateral motion artifact correction," *Front. Neuroinf.* **12**, 98 (2018).

103. R. Hattori and T. Komiyama, "Patchwarp: corrections of non-uniform image distortions in two-photon calcium imaging data by patchwork affine transformations," bioRxiv (2021).

104. S. Aghayee et al., "Particle tracking facilitates real time capable motion correction in 2d or 3d two-photon imaging of neuronal activity," *Front. Neural Circuits* **11**, 56 (2017).

105. M. Pachitariu, M. S. C. Carandini, and K. D. Harris, "Drift correction for electrophysiology and two-photon calcium imaging," *Comput. Syst. Neurosci.* III–82 (2018).

106. 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).

107. J. Voigts and M. T. Harnett, "An animal-actuated rotational head-fixation system for 2-photon imaging during 2-d navigation," bioRxiv 262543 (2018).

108. J. Voigts and M. T. Harnett, "Somatic and dendritic encoding of spatial variables in retrosplenial cortex differs during 2D navigation," *Neuron* **105**(2), 237–245.e4 (2020).

109. A. S. Charles et al., "Stochastic filtering of two-photon imaging using reweighted l_1," in *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, IEEE, pp. 1038–1042 (2017).

110. X. Li et al., "Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising," *Nat. Methods* **18**(11), 1395–1400 (2021).

111. J. Friedrich et al., "Multi-scale approaches for high-speed imaging and analysis of large neural populations," *PLoS Comput. Biol.* **13**(8), e1005685 (2017).

112. D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika* **81**(3), 425–455 (1994).

113. S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.* **9**(9), 1532–1546 (2000).

114. M. Weigert et al. "Content-aware image restoration: pushing the limits of fluorescence microscopy," *Nat. Methods* **15**(12), 1090–1097 (2018).

115. Z. Wei et al., "A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology," *PLoS Comput. Biol.* **16**(9), e1008198 (2020).

116. R. Nolan, "Algorithms for the correction of photobleaching," PhD thesis, University of Oxford (2018).

117. S. Kubler et al., "A robust and versatile framework to compare spike detection methods in calcium imaging of neuronal activity," in *IEEE 18th Int. Symp. Biomed. Imaging (ISBI)*, IEEE, pp. 375–379 (2021).

118. F. Cutrale et al., "Using enhanced number and brightness to measure protein oligomerization dynamics in live cells," *Nat. Protoc.* **14**(2), 616–638 (2019).

119. S. W. Keemink et al., "Fissa: a neuropil decontamination toolbox for calcium imaging signals," *Sci. Rep.* **8**(1), 3493 (2018).

120. H. Dana et al., "Thy1-GCaMP6 transgenic mice for neuronal population imaging in vivo," *PLoS One* **9**(9), e108697 (2014).

121. T. L. Daigle et al., "A suite of transgenic driver and reporter mouse lines with enhanced brain-cell-type targeting and functionality," *Cell* **174**(2), 465–480.e22 (2018).

122. F. Helmchen, "Calibration of fluorescent calcium indicators," *Cold Spring Harbor Protoc.* **2011**(8), pdb–top120 (2011).

123. A. Song et al., "Neural anatomy and optical microscopy (naomi) simulation for evaluating calcium imaging methods," *J. Neurosci. Methods* **358**, 109173 (2021).

124. M. Pachitariu, C. Stringer, and K. D Harris, "Robustness of spike deconvolution for neuronal calcium imaging," *J. Neurosci.* **38**(37), 7976–7985 (2018).

125. J. Friedrich, P. Zhou, and L. Paninski, "Fast online deconvolution of calcium imaging data," *PLoS Comput. Biol.* **13**(3), e1005423 (2017).

126. M. H. Evans, R. S. Petersen, and M. D. Humphries, "On the use of calcium deconvolution algorithms in practical contexts," bioRxiv 871137 (2019).

127. S. W. Jewell et al., "Fast nonconvex deconvolution of calcium imaging data," *Biostatistics* **21**(4), 709–726 (2020).

128. E. Pnevmatikakis and L. Paninski, "Sparse nonnegative deconvolution for compressive calcium imaging: algorithms and phase transitions," in *Adv. Neural Inf. Process. Syst.*, pp. 1250–1258 (2013).

129. R. Shibue and F. Komaki, "Deconvolution of calcium imaging data using marked point processes," *PLoS Comput. Biol.* **16**(3), e1007650 (2020).

130. L. Theis et al., "Benchmarking spike rate inference in population calcium imaging," *Neuron* **90**(3), 471–482 (2016).

131. P. Rupprecht et al., "A database and deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging," *Nat. Neurosci.* **24**(9), 1324–1337 (2021).

132. F. Helmchen and D. W. Tank, "A single-compartment model of calcium dynamics in nerve terminals and dendrites," *Cold Spring Harbor Protoc.* **2015**(2), pdb–top085910 (2015).

133. H. Lütcke et al., "Inference of neuronal network spike dynamics and topology from calcium imaging data," *Front. Neural Circuits* **7**, 201 (2013).

134. A. Speiser et al., "Fast amortized inference of neural activity from calcium imaging data with variational autoencoders," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, pp. 4027–4037 (2017).

135. J. T. Vogelstein et al., "Fast nonnegative deconvolution for spike train inference from population calcium imaging," *J. Neurophysiol.* **104**(6), 3691–3704 (2010).

136. J. Friedrich, P. Zhou, and L. Paninski, "Fast active set methods for online deconvolution of calcium imaging data," arXiv. org, Sept (2016).

137. L. Huang et al., "Relationship between simultaneously recorded spiking activity and fluorescence signal in GCaMP6 transgenic mice," *Elife* **10**, e51675 (2021).

138. E. Ganmor et al., "Direct estimation of firing rates from calcium imaging data," arXiv:1601.00364 (2016).

139. C. Li et al., "Fully affine invariant methods for cross-session registration of calcium imaging data," *eNeuro* **7**(4), 1–12 (2020).

140. S. Yadav et al., "Multi-session alignment for longitudinal calcium imaging," in *COSYNE* (2021).

141. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.* **60**(2), 91–110 (2004).

142. L. Sheintuch et al., "Tracking the same neurons across multiple days in ca2+ imaging data," *Cell Rep.* **21**(4), 1102–1115 (2017).

143. J. A. Cardin, M. C. Crair, and M. J. Higley, "Mesoscopic imaging: shining a wide light on large-scale neural dynamics," *Neuron* **108**(1), 33–43 (2020).

144. A. K. Churchland et al., "Dataset from: Single-trial neural dynamics are dominated by richly varied movements," http://repository.cshl.edu/id/eprint/38599 (2020).

145. C. P. Bengtson et al., "Nuclear calcium sensors reveal that repetition of trains of synaptic stimuli boosts nuclear calcium signaling in ca1 pyramidal neurons," *Biophys. J.* **99**(12), 4066–4077 (2010).

146. C. K. Kim et al., "Prolonged, brain-wide expression of nuclear-localized GCaMP3 for functional circuit mapping," *Front. Neural Circuits* **8**, 138 (2014).

147. J. V. Cramer et al., "In vivo widefield calcium imaging of the mouse cortex for analysis of network connectivity in health and brain disease," *Neuroimage* **199**, 570–584 (2019).

148. C. Ren and T. Komiyama, "Characterizing cortex-wide dynamics with wide-field calcium imaging," *J. Neurosci.* **41**(19), 4160–4168 (2021).

149. S. Musall et al., "Single-trial neural dynamics are dominated by richly varied movements," *Nat. Neurosci.* **22**(10), 1677–1686 (2019).

150. D. Barson et al., "Simultaneous mesoscopic and two-photon imaging of neuronal activity in cortical circuits," *Nat. Methods* **17**(1), 107–113 (2020).
151. L. Tian et al., "Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators," *Nat. Methods* **6**(12), 875–881 (2009).
152. S. Lohani et al., "Dual color mesoscopic imaging reveals spatiotemporally heterogeneous coordination of cholinergic and neocortical activity," bioRxiv (2020).
153. M. T. Valley et al., "Separation of hemodynamic signals from GCaMP fluorescence measured with wide-field imaging," *J. Neurophysiol.* **123**(1), 356–366 (2020).
154. S. W. Oh et al., "A mesoscale connectome of the mouse brain," *Nature* **508**(7495), 207–214 (2014).
155. G. Paxinos and K. B. J. Franklin, *The Mouse Brain in Stereotaxic Coordinates*, pp. 1–93, Elsevier (2001).
156. S. Romero et al., "Cellular and widefield imaging of sound frequency organization in primary and higher order fields of the mouse auditory cortex," *Cereb. Cortex* **30**(3), 1603–1622 (2020).
157. K. K. Sit and M. J. Goard, "Distributed and retinotopically asymmetric processing of coherent motion in mouse visual cortex," *Nat. Commun.* **11**(1), 1–14 (2020).
158. D. Zhi, M. King, and J. Diedrichsen, "Evaluating brain parcellations using the distance controlled boundary coefficient," bioRxiv (2021).
159. J. Liu et al., "Parallel processing of sound dynamics across mouse auditory cortex via spatially patterned thalamic inputs and distinct areal intracortical circuits," *Cell Rep.* **27**(3), 872–885.e7 (2019).
160. C. J. MacDowell and T. J. Buschman, "Low-dimensional spatiotemporal dynamics underlie cortex-wide neural activity," *Curr. Biol.* **30**(14), 2665–2680.e8 (2020).
161. M. Li et al., "Density center-based fast clustering of widefield fluorescence imaging of cortical mesoscale functional connectivity and relation to structural connectivity," *Neurophotonics* **6**(4), 045014 (2019).
162. E. M. R. Lake et al., "Spanning spatiotemporal scales with simultaneous mesoscopic Ca2+ imaging and functional MRI," bioRxiv 464305 (2018).
163. G. Mishne et al., "Calcium imaging data analysis (CIDAN): a multiscale approach for extraction of neuronal structures from calcium imaging data," in preparation.
164. E. M. R. Lake et al., "Simultaneous cortex-wide fluorescence ca 2+ imaging and whole-brain fMRI," *Nat. Methods* **17**(12), 1262–1271 (2020).
165. D. Xiao et al., "Mapping cortical mesoscopic networks of single spiking cortical or sub-cortical neurons," *Elife* **6**, e19976 (2017).
166. K. B. Clancy, I. Orsolic, and T. D. Mrsic-Flogel, "Locomotion-dependent remapping of distributed cortical networks," *Nat. Neurosci.* **22**(5), 778–786 (2019).
167. X. Liu et al., "Multimodal neural recordings with neuro-fitm uncover diverse patterns of cortical–hippocampal interactions," *Nat. Neurosci.* **24**(6), 886–896 (2021).
168. A. J. Peters et al., "Striatal activity topographically reflects cortical activity," *Nature* **591**(7850), 420–425 (2021).
169. W. Denk et al., "Imaging calcium dynamics in dendritic spines," *Curr. Opin. Neurobiol.* **6**(3), 372–378 (1996).
170. A. Kerlin et al., "Functional clustering of dendritic activity during decision-making," *Elife* **8**, e46966 (2019).
171. S. S. Suratkal, Y.-H. Yen, and J. Nishiyama, "Imaging dendritic spines: molecular organization and signaling for plasticity," *Curr. Opin. Neurobiol.* **67**, 66–74 (2021).
172. A. R. Graves et al., "Visualizing synaptic plasticity in vivo by large-scale imaging of endogenous ampa receptors," *Elife* **10**, e66809 (2021).
173. F. Ali and A. C. Kwan, "Interpreting in vivo calcium signals from neuronal cell bodies, axons, and dendrites: a review," *Neurophotonics* **7**(1), 011402 (2020).
174. B. L. Sabatini, M. Maravall, and K. Svoboda, "Ca2+ signaling in dendritic spines," *Curr. Opin. Neurobiol.* **11**(3), 349–356 (2001).
175. N.-L. Xu et al., "Nonlinear dendritic integration of sensory and motor input during an active sensing task," *Nature* **492**(7428), 247–251 (2012).

176. G. J. Broussard et al., "In vivo measurement of afferent activity with axon-specific calcium imaging," *Nat. Neurosci.* **21**(9), 1272–1280 (2018).

177. J. D. Seelig et al., "Two-photon calcium imaging from head-fixed drosophila during optomotor walking behavior," *Nat. Methods* **7**(7), 535–540 (2010).

178. N. Vajente et al., "Calcium imaging in drosophila melanogaster," *Adv. Exp. Med. Biol.* **1131**, 881–900 (2020).

179. R. Srinivasan et al., "Ca2+ signaling in astrocytes from ip3r2-/- mice in brain slices and during startle responses in vivo," *Nat. Neurosci.* **18**(5), 708–717 (2015).

180. Y. Wang et al., "Accurate quantification of astrocyte and neurotransmitter fluorescence dynamics for single-cell and population-level physiology," *Nat. Neurosci.* **22**(11), 1936–1944 (2019).

181. P. O'Herron et al., "Neural correlates of single-vessel haemodynamic responses in vivo," *Nature* **534**(7607), 378–382 (2016).

182. T. J. Ebner and G. Chen, "Use of voltage-sensitive dyes and optical recordings in the central nervous system," *Progr. Neurobiol.* **46**(5), 463–506 (1995).

183. M. H. Zhu et al., "Population imaging discrepancies between a genetically-encoded calcium indicator (GECI) versus a genetically-encoded voltage indicator (GEVI)," *Sci. Rep.* **11**(1), 1–15 (2021).

184. M. H. Mohajerani et al., "Mirrored bilateral slow-wave cortical activity within local circuits revealed by fast bihemispheric voltage-sensitive dye imaging in anesthetized and awake mice," *J. Neurosci.* **30**(10), 3745–3751 (2010).

185. T. Knöpfel and C. Song, "Optical voltage imaging in neurons: moving from technology development to practical tool," *Nat. Rev. Neurosci.* **20**(12), 719–727 (2019).

186. K. D. Piatkevich et al., "Population imaging of neural activity in awake behaving mice," *Nature* **574**(7778), 413–417 (2019).

187. M. E. Xie et al., "High-fidelity estimates of spikes and subthreshold waveforms from 1-photon voltage imaging in-vivo," *Cell Rep.* **35**(1), 108954 (2021).

188. J. N. Mait, G. W. Euliss, and R. A. Athale, "Computational imaging," *Adv. Opt. Photonics* **10**(2), 409–483 (2018)

189. L. Waller, "Physics-constrained computational imaging," *Proc. SPIE* **11469**, 114690M (2020).

190. D. Deb et al., "Programmable 3D snapshot microscopy with fourier convolutional networks," arXiv:2104.10611 (2021).

191. J. Wu et al., "Iterative tomography with digital adaptive optics permits hour-long intravital observation of 3D subcellular dynamics at millisecond scale," *Cell* **184**(12), 3318–3332.e17 (2021).

192. Y. Xue et al., "Single-shot 3d wide-field fluorescence imaging with a computational miniature mesoscope," *Sci. Adv.* **6**(43), eabb7508 (2020).

193. Y.-G. Yoon et al., "Sparse decomposition light-field microscopy for high speed imaging of neuronal activity," *Optica* **7**(10), 1457–1468 (2020).

194. Y. Zhang et al., "Computational optical sectioning with an incoherent multiscale scattering model for light-field microscopy," *Nat. Commun.* **12**(1), 1–11 (2021).

195. C. Moretti and S. Gigan, "Readout of fluorescence functional signals through highly scattering tissue," *Nat. Photonics* **14**(6), 361–364 (2020).

196. E. A. Pnevmatikakis, "Analysis pipelines for calcium imaging data," *Curr. Opin. Neurobiol.* **55**, 15–21 (2019).

197. P. Berens et al., "Standardizing and benchmarking data analysis for calcium imaging," in *COSYNE* (2017).

198. S. L. Smith and M. Häusser, "Parallel processing of visual space by neighboring neurons in mouse visual cortex," *Nat. Neurosci.* **13**(9), 1144–1149 (2010).

199. X. Cheng and G. Mishne, "Spectral embedding norm: Looking deep into the spectrum of the graph laplacian," *SIAM J. Imaging Sci.* **13**(2), 1015–1048 (2020).

200. S. Levy et al., "Cell-type-specific outcome representation in the primary motor cortex," *Neuron* **107**(5), 954–971.e9 (2020).

201. B. M. Li et al., "Calciumgan: a generative adversarial network model for synthesising realistic calcium imaging data of neuronal populations," arXiv:2009.02707 (2020).

202. G. Vanwalleghem, L. Constantin, and E. K. Scott, "Calcium imaging and the curse of negativity," *Front. Neural Circuits* **14**, 607391 (2021).
203. L. L. Glickfeld, M. H. Histed, and J. H. R. Maunsell, "Mouse primary visual cortex is used to detect both orientation and contrast changes," *J. Neurosci.* **33**(50), 19416–19422 (2013).
204. H. K. Kato, S. N. Gillet, and J. S. Isaacson, "Flexible sensory representations in auditory cortex driven by behavioral relevance," *Neuron* **88**(5), 1027–1039 (2015).
205. A. S. Charles et al., "Dethroning the fano factor: a flexible, model-based approach to partitioning neural variability," *Neural Comput.* **30**(4), 1012–1045 (2018).
206. A. Giovannucci et al., "Onacid: online analysis of calcium imaging data in real time," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, pp. 2378–2388 (2017).
207. A. Giovannucci et al., "Fiola: an accelerated pipeline for fluorescence imaging online analysis," (2021).
208. E. J. Botcherby, R. Juskaitis, and T. Wilson, "Scanning two photon fluorescence microscopy with extended depth of field," *Opt. Commun.* **268**(2), 253–260 (2006).
209. G. Theriault et al., "Extended two-photon microscopy in live samples with bessel beams: steadier focus, faster volume scans, and simpler stereoscopic imaging," *Front. Cell Neurosci.* **8**, 139 (2014).
210. W. Gobel, B. M. Kampa, and F. Helmchen, "Imaging cellular network dynamics in three dimensions using fast 3D laser scanning," *Nat. Methods* **4**(1), 73–79 (2007).
211. W. Yang et al., "Simultaneous multi-plane imaging of neural circuits," *Neuron* **89**(2), 269–284 (2016).
212. B. F. Grewe et al., "Fast two-layer two-photon imaging of neuronal cell populations using an electrically tunable lens," *Biomed. Opt. Express* **2**(7), 2035–2046 (2011).
213. G. Duemani Reddy et al., "Three-dimensional random access multiphoton microscopy for functional imaging of neuronal activity," *Nat. Neurosci.* **11**(6), 713–720 (2008).
214. "Allen institute brain observatory whitepaper," The Allen Institute, https://help.brain-map .org/display/observatory/Documentation (accessed 27 September 2019).
215. A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *Int. Interdisciplinary PhD Workshop (IIPhDW)*, IEEE, pp. 117–122 (2018).
216. J. Magland et al., "Spikeforest, reproducible web-facing ground-truth validation of automated neural spike sorters," *Elife* **9**, e55167 (2020).
217. A. S. Charles et al., "Toward community-driven big open brain science: open big data and tools for structure, function, and genetics," *Annu. Rev. Neurosci.* **43**, 441–464 (2020).
218. S. A. Romano et al., "An integrated calcium imaging processing toolbox for the analysis of neuronal population dynamics," *PLoS Comput. Biol.* **13**(6), e1005526 (2017).
219. T. Abe et al., "Neuroscience cloud analysis as a service," bioRxiv 2020–2026 (2021).

Hadas Benisty is an associate research scientist in the Neuroscience Department at Yale University. She is an expert in developing interpretable models, aiming to investigate the dynamics of high-dimensional neuronal networks and how their organizational principles and plasticity relate to behavior. Before joining Yale University, she was a post-doc in the EE Department at Technion. She received her PhD in electrical engineering and BSc degrees in electrical engineering and physics from Technion.

Alexander Song received his BS degree from Cornell University in 2013 and his MS and PhD degrees from Princeton University in 2019. His doctoral research focused on developing microscopy techniques for neuroscience research. He joined as a postdoctoral researcher at MPI for Intelligent Systems, Stuttgart, in 2019, and is working on developing hardware for optical computation.

Gal Mishne is an assistant professor in the Halıcıoğlu Data Science Institute (HDSI) at UCSD, and affiliated with the ECE and CSE Departments and the Neurosciences Graduate Program. Before joining UCSD, she was a Gibbs assistant professor in Applied Math at Yale University, with Prof. Ronald Coifman's research group. She received her PhD in electrical engineering from Technion in 2017. Her research interests include high-dimensional data analysis, manifold learning, and computational neuroscience.

Adam S. Charles is an assistant professor in the Department of Biomedical Engineering at Johns Hopkins University. He received his master's and bachelor's degrees from Cooper Union in NYC and his PhD in electrical and computer engineering from Georgia Tech. He then held a post-doc position at the Princeton Neuroscience Institute. His interests are at the intersection of computational and theoretical neuroscience and data science as well as focusing on next-generation algorithms for extracting meaning from complex neural data.