Suite3D: Volumetric cell detection for two-photon microscopy

Ali Haydaroğlu[1], Sam Dodgson[1], Michael Krumin[1], Andrew Landau[1], Liad J. Baruchin[2], Tinya Chang[1], Jingkun Guo[3], David Meyer[3], Charu Bai Reddy[1], Jian Zhong[4], Na Ji[4], Sylvia Schröder[2], Kenneth D Harris[1], Alipasha Vaziri[3], Matteo Carandini[1]

In two-photon imaging of neuronal activity it is common to acquire 3-dimensional volumes. However, these volumes are typically processed plane by plane, leading to duplicated cells across planes, reduced signal-to-noise ratio per cell, uncorrected axial movement, and missed cells. To overcome these limitations, we introduce Suite3D, a volumetric cell detection pipeline. Suite3D corrects for 3D brain motion, estimating axial motion and improving estimates of lateral motion. It detects neurons using 3D correlation, which improves the signal-to-background ratio and detectability of cells. Finally, it performs 3D segmentation, detecting cells across imaging planes. We validated Suite3D with data from conventional multi-plane microscopes and advanced volumetric microscopes, at various resolutions and in various brain regions. Suite3D successfully detected cells appearing on multiple imaging planes, improving cell detectability and signal quality, avoiding duplications, and running >20x faster than a prior volumetric pipeline. Suite3D offers a powerful solution for analyzing volumetric two-photon data.

Introduction

The wide adoption of 2-photon microscopes has made it common to image 3D volumes of tissue, greatly increasing the number of cells that are sampled in the imaged region. With conventional 2-photon microscopes, volumetric imaging is typically achieved by scanning each plane in sequence[1]. This approach has yielded recordings from increasing numbers of neurons, from over 10^3 neurons in a 0.03 mm^3 volume[2] to over 10^4 neurons in a 0.3 mm^3 volume[3]. These numbers can be further increased with imaging systems that use temporal multiplexing approaches to obtain volumetric imaging. One such approach is Free-space Angular-Chirp-Enhanced Delay (FACED) imaging, which turns a light pulse into an ultrafast linear sequence of pulses, enabling passive line scanning[4–6]. Another approach is Light Beads Microscopy (LBM), which turns a light pulse into a column of light beads at multiple depths, thus acquiring an entire volume while scanning only one plane[7]. This approach has the potential of recording up to 10^6 neurons in a 15 mm^3 volume[7,8]. Beyond 2-photon imaging, methods such as SCAPE are capable of fast, high-resolution volumetric imaging of *in vivo* neural activity[9,10].

These approaches to volumetric imaging produce vast datasets, which are typically processed one plane at a time or segmented manually[2,5]. Processing 2-photon data requires a pipeline that performs a sequence of operations from motion correction to cell segmentation. Established pipelines such as Suite2p[11,12], CaImAn[13] and EXTRACT[14] perform this sequence of operations independently on individual planes. They are typically used even if the data are volumetric[3,8], sometimes with post-hoc corrections to merge cells that appear in multiple imaging planes[7,15]. By contrast, efforts to extend the processing to volumes, such as volumetric CaImAn[13], have been met by limited adoption perhaps due to their extensive computational demands.

[1]University College London, London, UK; [2]University of Sussex, Brighton, UK; [3]Rockefeller University, New York, NY, USA; [4]University of California Berkeley, CA, USA

When applied to volumetric data, however, plane-by-plane methods can lead to cell duplications, signal losses, and missed neurons. In the mouse cortex, for instance, neurons have somas that easily exceed ~20 µm in diameter[16,17] and pyramidal neurons may extend more in the axial dimension due to their apical dendrites. Moreover, the 2-photon point spread function is much more extended in the axial than lateral dimension, typically on the order of 10 µm[18]. Together, these factors mean that fluorescence from single neurons can cross multiple imaging planes even if these planes are ≥30 µm apart. Analyzing the planes one at a time therefore leads to multiple problems (Figure 1a-c). First, if a cell has significant signal on two or more planes, it will be split and detected independently in each of those planes, causing duplications. Second, when a cell is split across planes, the region detected on each plane will have fewer pixels and therefore lower signal than the full cell. Third, a cell may be below the detection threshold on one (or all) individual planes, and only detectable when planes are analyzed together. Finally, any brain movement or drift perpendicular to the imaging plane will be uncorrected, and its effects potentially mistaken for functional signals. These concerns are not just hypothetical. For instance, even in conventional 2-photon imaging with 20 µm inter-plane spacing, neurons clearly straddle multiple planes (Supplementary Figure 1).

To overcome these limitations, we developed Suite3D, a fully volumetric cell detection pipeline built for large-scale volumetric recordings. Suite3D extends to 3 dimensions the 2D pipeline of Suite2p, with novel 3D algorithms, computational improvements, and visualization tools. It is faster than plane-by-plane methods, and runs dramatically faster than volumetric CaImAn. It resolves the problems in plane-by-plane processing by performing all processing in 3D: first 3D motion correction (Figure 1g), then 3D detection (Figure 1h) and finally 3D segmentation (Figure 1i). The algorithms in Suite3D are robust to variations in signal levels and are tunable, with intuitive parameter selection and with visualization interfaces for semi-automated curation.

We validated Suite3D with data acquired with conventional 2-photon microscopes and with dedicated volumetric microscopes. Suite3D successfully detected cells appearing on multiple imaging planes, improving their detectability, signal quality, and identification. By contrast, plane-by-plane segmentation led to duplication of cells and to reduced SNR. Suite3D is thus a promising tool for analyzing 2-photon volumes acquired with a range of microscopy approaches.

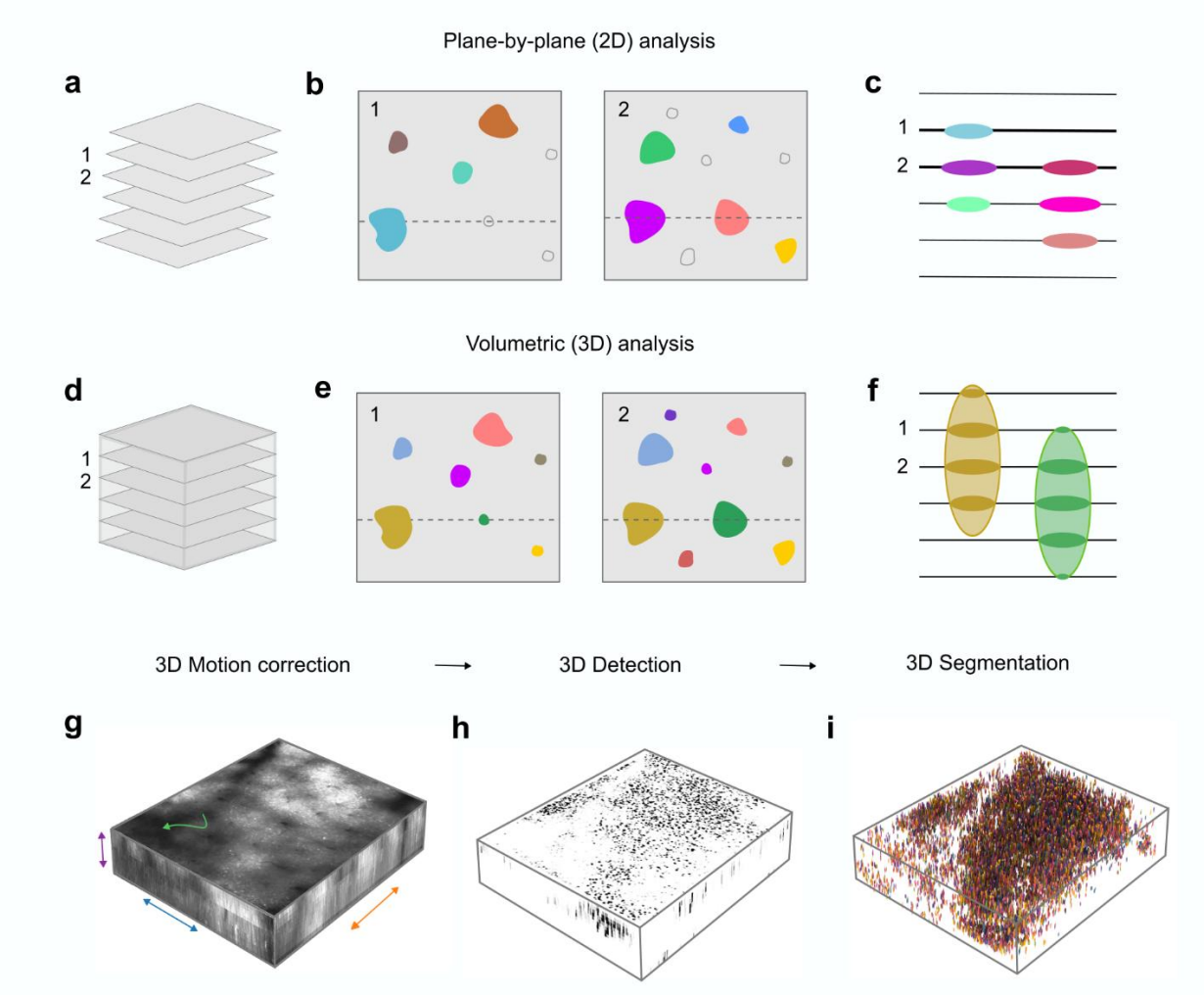


Figure 1**. Volumetric cell detection with Suite3D and large-scale 2-photon imaging. a**, A stack of imaging planes analyzed plane-by-plane by a 2D pipeline. **b**, Cartoon of two neighboring planes where colored cells are detected by the 2D pipeline, and undetected cells are outlined in grey. **c**, Cross section of the volume through the dashed line in (b), showing two cells that are potentially split into multiple planes. **d-f**, The same stack of imaging planes, analyzed by Suite3D as a volume. Detected cells extend across multiple planes. **g**, The first stage in Suite3D is 3D motion correction, which corrects for rigid motion in all three axes, and for non-rigid motion in the plane. **h**, The second stage is 3D detection, which produces a correlation volume. **i**, The third stage is 3D segmentation, which assigns voxels to 3D regions of interest (ROIs)**.**

Results

Suite3D comprises three core processing steps. The first is 3D motion correction, which corrects for lateral and axial brain movements (Figure 1g). The second is 3D detection, through spatiotemporal filtering, normalization, and thresholding (Figure 1h). The third is 3D segmentation, where voxels are assigned to distinct regions of interest (ROIs, Figure 1i). Below we describe these steps in order.

3D Motion Correction

The first stage in the Suite3D pipeline corrects for brain motion, computing this correction volumetrically rather than plane-by-plane as in standard pipelines. By correcting motion volumetrically, Suite3D can

estimate lateral shifts that are coherent across planes. Moreover, it can estimate axial shifts, which would be undetectable in individual planes.

The volumetric motion correction is based on phase correlation. We took the basic idea for motion correction in Suite2p[11] and extended it to 3D, and GPU-accelerated it to make it tractable. At each time step, Suite3D computes the 3D Fourier transform of the acquired volume and its 3D phase correlation with a reference volume. From the 3D phase correlation, it then computes x-, y- and z-shifts for each frame, which are applied to the frame in image space to correct motion. This volumetric motion correction is rigid and is followed by an optional GPU-accelerated non-rigid motion correction step applied per-plane, which corrects movements for each sub-region of each plane.

This 3D motion correction reduces noise and improves estimates of lateral brain motion compared to plane-by-plane methods. To compare the two methods, we analyzed the same volume of LBM data with Suite3D and Suite2p. The plane-by-plane motion correction implemented by Suite2p often estimated different shifts even for nearby planes (Figure 2a). This disparity across planes is unlikely to reflect genuine shearing of brain tissue, as the planes in this volume are spaced only 20 μm apart and are acquired essentially simultaneously (the light beads that define the planes are only 8 ns apart). Instead, the variance is due to the low amplitude of the phase correlation maps at each timepoint (Figure 2b,c) which make it difficult to estimate their peaks in the presence of noise. By contrast, the 3D motion correction implemented by Suite3D produced more robust motion estimates, and it did so in three dimensions instead of two (Figure 2d), with much sharper phase correlation maps (Figure 2e,f).

Moreover, Suite3D accurately estimates and corrects for axial motion. To verify the ability of Suite3D to estimate axial motion estimation, we performed recordings where we introduced intentional axial motion by moving the objective while recording. We could then ask whether Suite3D is able to reconstruct and correct this axial motion. Suite3D was able to accurately correct for the ground-truth axial motion up to a range of 40 μm in a 7-plane recording spanning 120 μm (Figure 2g). By contrast, 2D detection pipelines cannot detect axial motion, but rather suffer from its consequences. Indeed, axial motion can lead to cells shifting in-and-out-of plane. This shifting will lower the signal/noise ratio, and if it is systematic (e.g. if it correlates with some behavior), it may even distort the functional characterization of cells.

To perform these operations efficiently, Suite3D exploits “pipelining” and GPU acceleration. Volumetric microscopy produces large data sets - for instance, the LBM produces >700 GB/hour. Therefore, it is often unfeasible to store raw data locally on the processing computer, and data loading becomes a major speed bottleneck. Suite3D substantially reduces this problem by pipelining: it uses a separate thread to load the next batch from disk while the current batch is being processed. In addition, the Fourier transform and phase correlation operations are accelerated on a GPU with custom CUDA kernels.

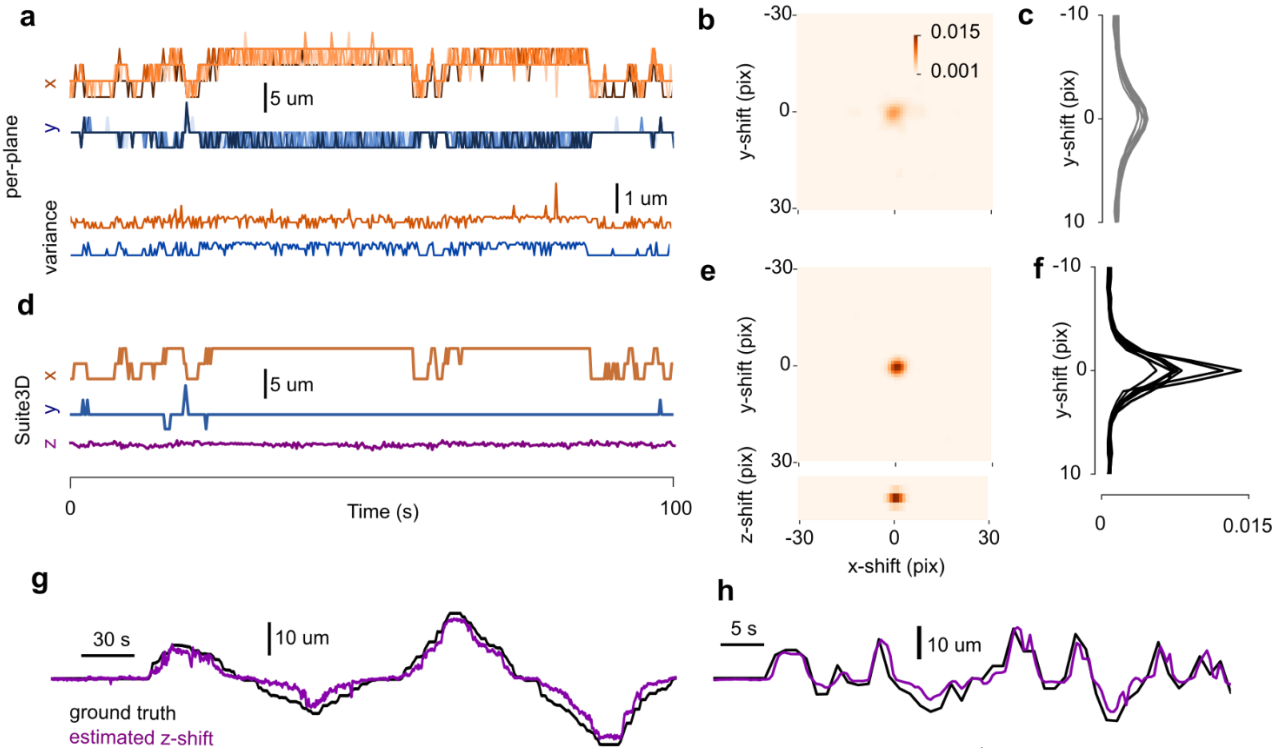


Figure 2. **3D motion correction predicts axial motion and improves planar shift estimates**. **a**, Standard 2D motion correction estimates x and y shifts independently for each plane (overlaid), leading to substantial variance of the estimated shifts across planes (bottom). **b**, 2D phase correlation map for one plane at a single timepoint in (a). **c.** Cross section of the phase correlation map in (b) for 10 randomly selected frames. **d**, Suite3D estimates a single set of x, y, and z shifts for the whole volume (shifts in z are sub-pixel). **e**. 3D phase correlation map for a single timepoint. **f**, Cross sections of the phase correlation map in (e), for the same 10 randomly selected frames in (c). **g**, In a recording with imposed axial drift and oscillations, the estimated sub-pixel z-shift (purple) matches the ground-truth objective movement (black).

3D Detection

The next stage in the Suite3D pipeline is to detect regions of interest (ROIs) through spatiotemporal processing steps that convert the (4D) movie – a sequence of volumes – into a 3D correlation volume. This volume highlights pixels that have high correlation to their neighbors at cell-like spatiotemporal scales (or at smaller scales to image synapses, as in Figure 5b). The procedure is related to the 2D correlation map in Suite2p[12], adapted to 3D and modified to be computationally tractable on large-scale datasets. The motion-corrected movie (Figure 3a) is high-pass filtered in time (typically with a 0.1 Hz cutoff for slow calcium indicators) and each pixel is normalized by the standard deviation of its first temporal differences (Figure 3b) – these steps exclude slow temporal drifts and normalize for differences in signal across the FOV. The resulting movie is then band-pass filtered in space (typically between 5 and 70 μm) to subtract neuropil and highlight cell-sized structures. To reduce this filtered movie to a single 3D volume, one could take the mean or root-mean-square of each pixel over time; however, if a cell were to only activate a few times in a long recording, it would be hard to detect. Instead, each pixel in the movie is thresholded before taking the root-sum-square (Figure 3c). To make these steps feasible on extremely large datasets (several TB), Suite3D parallelizes and memory-maps them. Suite3D also includes an intuitive parameter-sweeping interface that allows the user to systematically tune parameters and improve the correlation volume.

Computing this 3D correlation volume improves the signal-to-background ratio and detectability of cells compared to 2D methods. Computing 2D correlation maps for each plane and stitching them together into a volume throws out axial correlations that are present in volumetric data (Supplementary Figure 2). When applied to the same dataset, the 2D correlation map (Figure 3f) has substantially more background noise than the 3D correlation volume (Figure 3c). Moreover, cells with small footprints on a given plane (e.g., Figure 3c, arrows) would not be visible in a 2D correlation map (Figure 3f, arrows). Rather than contributing to the signal of the imaged neuron, those pixels would contribute to the background noise.

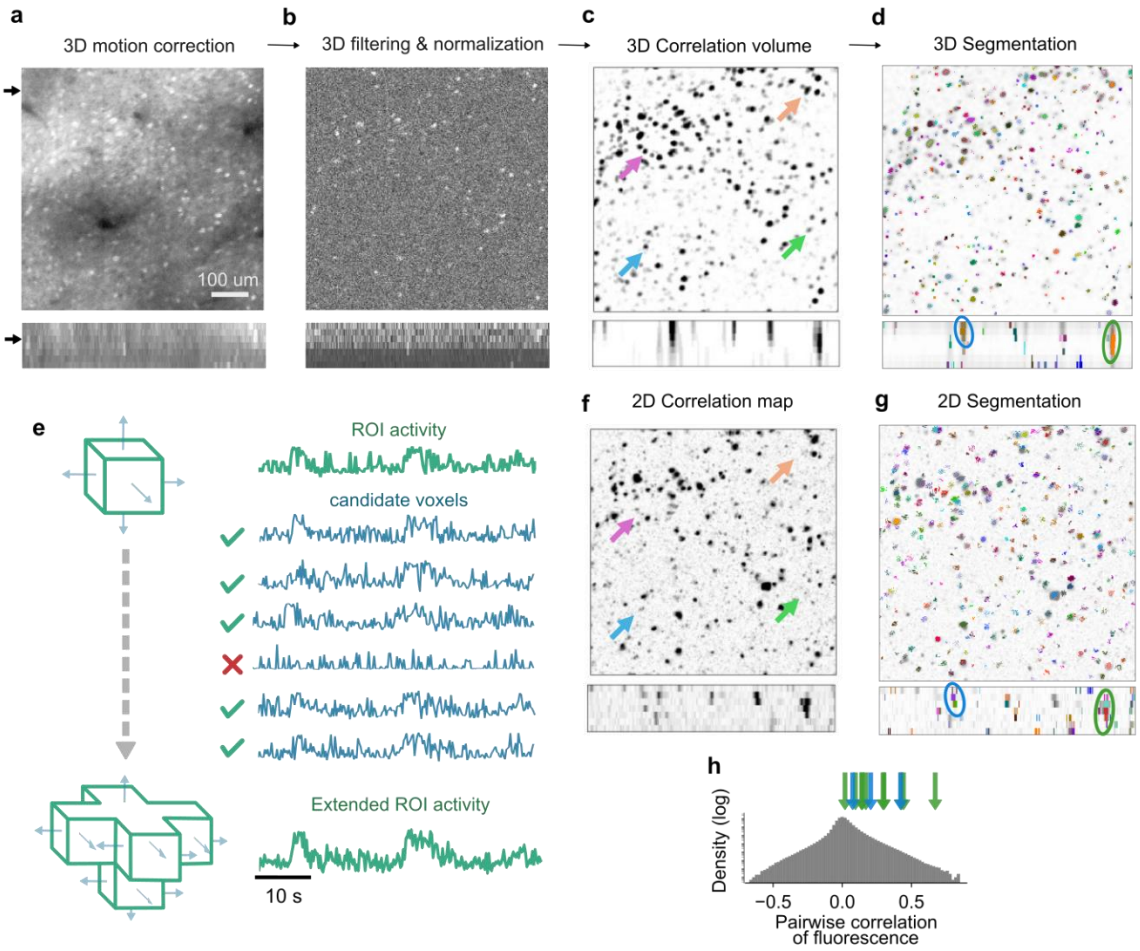


Figure 3. **3D correlation and segmentation improve cell detectability and avoid double-counted cells**. **a**, The motion-corrected 3D movie, seen after standard-deviation projection across time in an example plane (*top*) and an example vertical section (*bottom*). **b**, The temporally high-pass filtered and pixel-wise normalized movie (standard-deviation projection). **c**, The correlation volume, computed by spatial band-passing, pixel-wise temporal thresholding and standard-deviation projection. **d**, The segmented ROIs (pre-curation) overlaid on the correlation volume. **e**, The segmentation algorithm starts with a single seed voxel from the correlation map and computes the correlation of its activity to neighboring candidate voxels on active frames. Candidates that pass a correlation threshold join the seed voxel to form the new ROI, which is iteratively expanded via the same mechanism. **f**, Correlation map computed plane-by-plane. Colored arrows indicate cells visible in the 3D map and not in the plane-by-plane map. **g**, Results of plane-by-plane 2D segmentation. **h**. Histogram of pairwise correlations of all pairs of ROIs detected with 2D segmentation. Arrows indicate the pairwise correlations between 2D ROIs from different planes for two example cells in (d) and (g).

3D Segmentation

The final stage in the Suite3D pipeline is 3D segmentation. Similar to the 2D algorithm in Suite2p[12], Suite3D performs segmentation as a nested iterative process. Peak voxels of the correlation volume are expanded by including neighboring voxels that pass a correlation threshold in the normalized, neuropil-subtracted movie (Figure 3e). However, previous pipelines require the entire movie to be loaded into RAM at once, which is unrealistic for volumetric movies obtained with a large field of view or long duration. Suite3D reduces the RAM requirement to typical levels (<64 GB) while maintaining processing speed by parallelizing across ROIs and movie patches. Extremely large movies (~ 1TB) which would otherwise be impossible to segment can be processed with Suite3D on a typical workstation. The volumetric implementation of the algorithm introduces several interpretable parameters and a tuning interface for optimizing the pipeline for challenging datasets (e.g. high-noise, low-resolution).

3D segmentation detects cells that span multiple imaging planes, which would be duplicated with two-dimensional segmentation methods (Figure 3d,g). Although post-hoc merging could potentially overcome this duplication, it is not trivial to merge 2D ROIs across planes. One implementation of post-hoc merging uses pairwise correlations of ROIs with nearby xy-coordinates on neighboring planes, and merges them if these correlations are above a threshold[7]. However, we show that the pairwise correlations of ROIs corresponding to the same cell can be very low, especially if an ROI has a small footprint on a given plane. The pairwise correlation of activity in ROIs belonging to the same cell falls well within the distribution of pairwise correlations of all ROIs within a recording, and a hard threshold will not sufficiently merge ROIs (Figure 3h,i). Thus, using plane-by-plane segmentation methods in any multi-plane recording with dense spacing (<30 μm) can lead to over-counting of cells.

After segmentation, a time course of neural activity is obtained for each ROI by performing a weighted average, neuropil subtraction, and deconvolution. The mask of each ROI is used to compute a weighted average of the fluorescence of each ROI pixel from the motion-corrected movie. The neuropil fluorescence is then calculated in a hollow shell around each ROI (excluding nearby cells) and is subtracted from the ROI fluorescence. The neuropil-subtracted fluorescence is then deconvolved to counteract the effects of the indicator's time course[19,20].

Validation

In describing the three stages of processing of Suite3D, we have highlighted some of its advantages over 2D methods in improving motion correction, cell detectability, and cell identification. Its 3D motion correction reduces noise and improves estimates of lateral brain motion, and accurately estimates and corrects for axial motion (Figure 2). Its 3D ROI detection improves the signal-to-background ratio and detectability of cells compared to 2D methods (Figure 3c,f,h). Its 3D segmentation detects cells that span multiple imaging planes, which would otherwise be duplicated (Figure 3).

To quantify the advantages of Suite3D, we compared it to Suite2p and found that it improves the signal-to-noise ratio for cells detected in both pipelines, and detects a larger number of stimulus-responsive cells. We presented multiple repeats of drifting grating stimuli to mice while imaging the visual cortex with an LBM. We analyzed the resulting dataset with default parameters on both pipelines and compared the raw fluorescence traces and deconvolved traces for ROIs detected by each pipeline. First, we compared cells that were identified in both pipelines. The dF/F traces for matched cells that were extracted by Suite3D were smoother and appeared to have less shot noise (Figure 4a). A quantitative comparison of shot noise level per cell using an estimator from ref.[21] confirmed that the Suite3D-extracted traces for matched cells were less noisy (Figure 4b). To confirm this corresponds to an increase in functional signal-to-noise ratio, we compared the standard deviation of dF/F across multiple repeats

of the same stimulus for each cell, and found that stimulus responses were more reliable in Suite3D-extracted cells (Figure 4c). Next we asked whether there were cells that were only detectable with Suite3D and below the threshold of detection in plane-by-plane methods (Figure 1b,e). We considered only ROIs that were significantly responsive to visual stimuli (see Methods). We found that Suite3D identified a larger number of stimulus-responsive cells in the same dataset (Figure 4d,e). Suite3D discovered a larger number of cells that had a lower fraction of stimulus-related variance, indicating that these may have been below the threshold of detection in Suite2p.

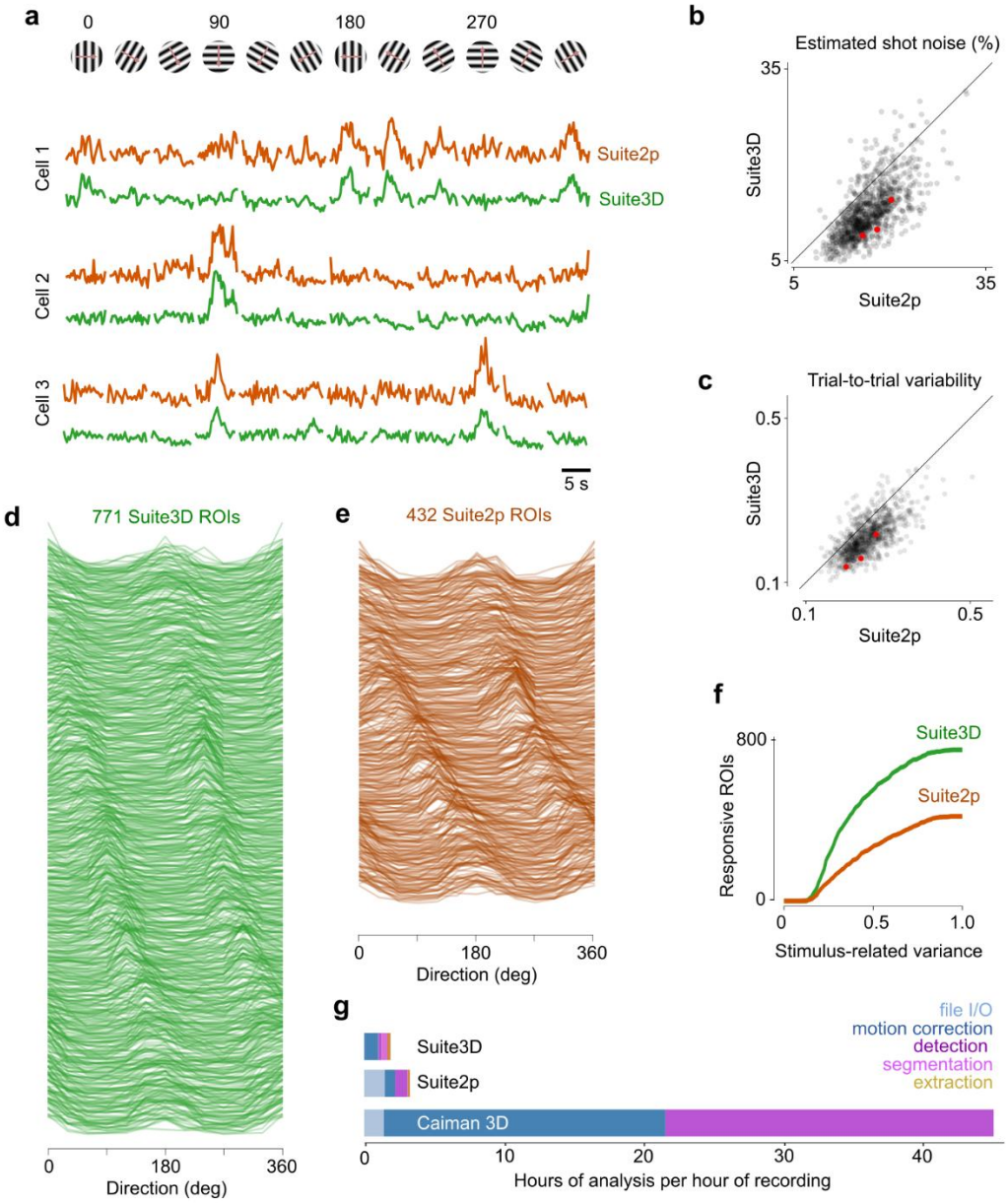


Figure 4. **Suite3D ROIs are less noisy and more numerous than Suite2p ROIs. a**, dF/F traces of three example ROIs in visual cortexidentified by both Suite2p and Suite3D in response to a single repeat of drifting gratings. **b**, For ROIs matched across the two pipelines, the estimated fraction of shot noise in the extracted fluorescence [21]. Red dots indicate the cells shown in (a). **c**, Standard deviation of stimulus responses across 8 repeats for ROIs matched across the two pipelines. **d**, Direction tuning curves for all significantly stimulus-responsive ROIs found by Suite3D in an example session. **e**, Same as (d), for Suite2p ROIs. **f**, Cumulative histogram of stimulus-related variance for all stimulus-responsive ROIs found by each pipeline, computed on deconvolved traces. **g**, Runtime of Suite2p, Suite3D and volumetric CaImAn per hour of recording.

We also compared Suite3D to a prior 3D analysis pipeline, the 3D version of CaImAn[13], and found that it was 24 times faster. To make the comparison, we considered a volume with 20 x 933 x 733 voxels, acquired at a 4.1 Hz rate with a Light Beads Microscope (LBM). Running the 3D version of CaImAn on this recording required 45.3 minutes of processing time per minute of recording, whereas plane-by-plane Suite2p required 3.3 minutes and Suite3D required 1.9 minutes (Figure 4g). While Suite2p and Suite3D were able to process the full 75 minute recording, CaImAn 3D was unable to process the full recording due to insufficient memory on a workstation with 512 GB of RAM, and therefore was run on a 5-minute subset. The difficulty of running the 3D version of CaImAn on large datasets made it unfeasible to perform further comparisons with Suite3D.

Table 1. Datasets used in the development and testing of Suite3D.

Lab	Microscope	Imaging target	Indicator	FOV (mm)	Axial spacing	Plane acquisition
UCL	LBM	Visual cortex L2/3-L4	Transgenic GCamp6s	4.0 x 4.0 x 0.4	20 µm	Simultaneous
RU	LBM	Visual cortex L2/3-L4	Transgenic GCamp6s	4.0 x 4.0 x 0.4	15 µm	Simultaneous
UCL	Conventional multi-plane 2P	Hippocampus CA1	Viral GCamp6f	0.7 x 0.7 x 0.1	20 µm	Sequential
UCL	Conventional multi-plane 2P	Visual cortex L5	Transgenic GCamp6s	0.7 x 0.7 x 0.05	10 µm	Sequential
UCL	Conventional multi-plane 2P	Visual cortex L5	Transgenic GCamp6s	0.7 x 0.7 x 0.15	20 µm	Sequential
UCB	FACED	Cortex L2/3-L4	Transgenic GCamp6s	1.0 x 1.0 x 0.2	5 µm	Sequential
US	High-resolution multi-plane 2P	Superior colliculus (retinal boutons)	Viral SyGCaMP8m	0.14 x 0.14 x 0.01	2 µm	Sequential

Finally, we validated Suite3D over a range of imaging modalities, resolutions and brain regions. We have seen that Suite3D provides high-quality results on volumetric data where planes are acquired simultaneously as in Light Beads Microscopy (Figure 2, Figure 3). Suite3D is also well-suited for more common imaging configurations where planes are acquired sequentially, such as conventional multi-plane 2-photon imaging with a piezo-driven objective. In mouse hippocampus and cortex we find that inter-plane spacings of 20 µm lead to pyramidal neurons crossing multiple imaging layers, as illustrated in the example dataset of dense excitatory neurons in CA1 (Figure 5b). To verify the performance of Suite3D on higher resolution recordings of non-somatic targets, we tested it on recordings of retinal boutons in superior colliculus (Figure 5c). We also applied Suite3D to volumetric recordings acquired with the novel FACED imaging method that enables ultrafast line-scanning (Figure 5d)[4–6].

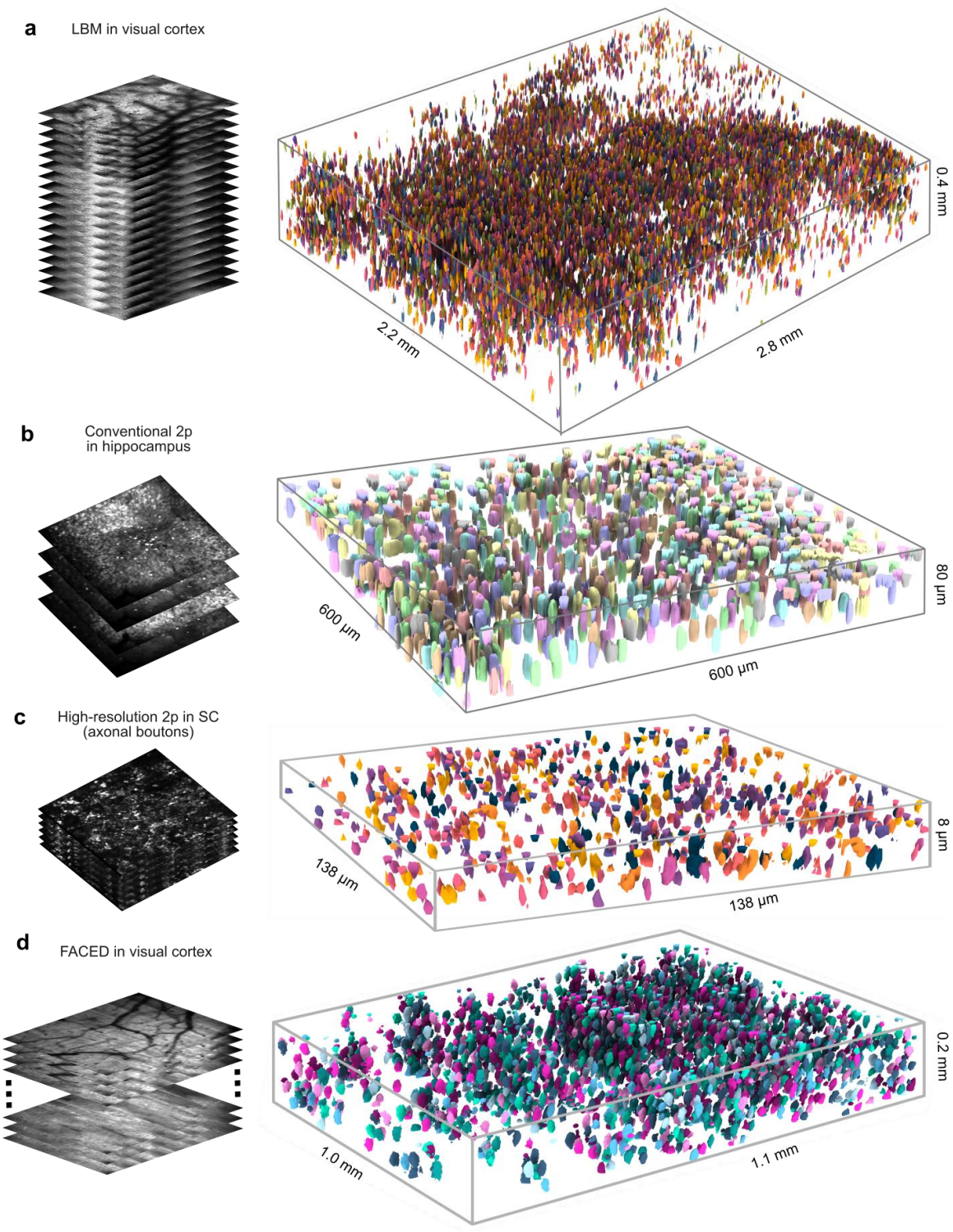


Figure 5. **Suite3D works on volumes imaged across modalities, scales, noise levels and resolutions. a**, A volume imaged with Light Beads Microscopy (LBM) in mouse visual cortex, with dense expression of GCaMP6s in excitatory neurons (*left*). The >20,000 neurons identified by Suite3D (*right*). **b**, Same, for a volume acquired with conventional multi-plane 2-photon imaging, with sequentially-acquired planes in mouse hippocampus with dense excitatory expression of GCaMP6s. **c**, Same, for a volume acquired with multi-plane 2-photon imaging at high resolution in mouse Superior Colliculus, showing boutons on retinal axons. **d**, Same for a volume imaged with FACED in mouse visual cortex with dense expression of GCaMP6s in excitatory neurons.

While Suite3D performed well with default parameters, its performance could be further optimized with built-in parameter tuning interfaces. A full list of testing datasets is listed in Table 1, and they span a range of sizes, resolutions and targets recorded across four laboratories. Recommended parameters of Suite3D perform well across datasets; however, as is the case for all processing pipelines, results sometimes benefit from user-guided optimization of parameters. This can be difficult in other pipelines: it may require re-running the entire dataset many times, writing code to extract intermediate results, or creating custom interfaces to visually compare segmentation results across tens of runs. Suite3D provides parameter sweeping interfaces that are integrated into the main processing pipeline, can sweep tens of combinations of parameters on subsets of data in minutes in the detection and segmentation steps, and visualize intermediate results for the user to directly compare and gain an intuition of the parameters. A robust set of default parameters with this optimization interface make Suite3D well suited for accurately and efficiently analyzing volumetric data obtained with a variety of 2-photon imaging methods.

Discussion

Volumetric datasets are increasingly common in 2-photon microscopy, and Suite3D is ideally positioned to process them because it takes into consideration aspects of the data that are intrinsically 3-dimensional. First, brain motion is pervasive and is likely to be correlated across planes, and to occur in all directions including the axial one. Second, unless the planes are separated by distances that are substantially larger than the structures of interest (typically, neurons), those structures will appear in multiple planes. Indeed, depending on the target region and axial point spread function of the imaging system, the images of neurons can extend >30 μm axially. Third, the main source of structured noise in 2-photon imaging is neuropil contamination, and this source is three-dimensional[22].

Suite3D takes these aspects of the data in consideration by performing all its operations in 3D. First, it performs motion correction in 3D and estimates not only horizontal motion but also axial motion. Second, Suite3D detects neurons using 3D correlation, which improves the signal-to-background ratio and detectability of cells. Finally, it performs 3D segmentation, detecting cells across imaging planes.

Suite3D is thus a major improvement over current standard methods which process the data in 2D. Typical 2D analysis pipelines can only correct for horizontal brain motion. This can lead to artifactual signals when motion in the axial direction leads to cells coming in and out of the imaging plane. Moreover, analyzing the data plane by plane leads to duplicated cells across planes, reduced signal-to-noise ratio per cell, and missed cells. Duplication could be reduced with manual or semi-automated curation [23,24] but such approaches are not scalable to larger-scale recordings and are not validated by widely accepted criteria for merging of duplicates.

Suite3D improves signal-to-noise ratio over plane-by-plane methods across each processing step in the pipeline, largely driven by a simple principle: more voxels means better signal. Estimates of brain motion can be noisy in single planar frames, and combining information across a volumetric frame allows better estimates. Including spatial correlations in the axial direction improves the detectability of cells and allows more robust segmentation. Detected cells have more voxels in Suite3D as they are no longer split over multiple planes, therefore extracted cells have lower levels of shot noise. Indeed, we demonstrated that Suite3D detects more cells with stronger neural responses compared to plane-by-plane methods.

We have shown that Suite3D operates successfully on volumetric data from multiple microscopes, resolutions and brain regions. In all these datasets, it successfully detected cells appearing on multiple imaging planes, improving cell detectability and signal quality, and avoiding duplications. This last

feature alone – avoiding duplications – should arguably be a reason to switch to Suite3D when analyzing volumetric data from any kind of 2-photon microscope. Indeed, we have shown that even conventional 2-photon microscopy with planes separated by 20 um leads to multiple cells appearing on multiple planes (Supplementary Figure 1). This is a common imaging configuration, and any lab that adopts it might benefit from analyzing the data with Suite3D. Moreover, volumetric imaging is becoming even more popular thanks to volumetric 2-photon microscopes such as FACED[4–6] and LBM[7,8], which image multiple planes in quick succession. As we have shown, data from these microscopes is ideally suited to analysis with Suite3D.

The user experience of Suite3D is driven by two design principles: processing should be fast, and its steps should be transparent. Speed is an obvious concern when dealing with methods such as FACED[4–6] or LBM[7,8], which generate close to 1 TB of data per hour, but it also has a less obvious benefit: faster algorithms allow advanced users to tinker and optimize parameters on challenging datasets. Otherwise, if a pipeline takes many hours to run, users are likely to run it only once as a “black box” and accept the results as they are, without fine tuning the parameters. By running quickly, Suite3D allows users to iterate. It allows them to inspect intermediate results and easily optimize parameters through a graphical interface, so that they can turn each knob and understand what it does.

Suite3D is thus a useful tool for volumetric imaging applicable across a variety of microscopes and scientific questions. We released it to the community as an open-source package (github.com/alihaydaroglu/suite3d) and welcome suggestions and contributions to further improve it.

Acknowledgments

We thank Carsen Stringer and Marius Pachitariu for invaluable suggestions at the beginning of this project and their continued development of Suite2p. We thank Bex Terry, Kevin Barber, Francisca Martínez Traub, and the UCL and Rockefeller University animal facilities for animal care and management; Tobias Nöbauer and Jeffrey Demas for advising microscope development; Santiago Otero Colonel, Flynn O'Connell, Paul Fahey and Max Gagnon for testing the software; and Kimberly Ren for feedback on the manuscript and figures. This work was funded by UKRI (Frontier Award EP/X022366/1 to MC), BBSRC (grant BB/W019884/1 to MC), the National Institutes of Health BRAIN initiative (grant U01NS126057 to AV and MC), the Wellcome Trust (Investigator Award 223144/Z/21/Z to MC and KDH), and the ERC (101097874 to KDH), the National Institutes of Health BRAIN initiative (UF1NS107696, U01NS118300, U01NS137449 to NJ and JZ) , and jointly by the Wellcome Trust and the Royal Society (Sir Henry Dale Fellowship 220169/Z/20/Z to SS). 3D visualizations were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. MC holds the GlaxoSmithKline / Fight for Sight Chair in Visual Neuroscience.

Author contributions

According to the CRediT taxonomy.

	Bai Reddy	Baruchin	Carandini	Chang	Dodgson	Guo	Harris	Haydaroğlu	Ji	Krumin	Landau	Meyer	Schröder	Vaziri	Zhong
Conceptualization			●				●	●							
Data curation								●							
Formal analysis								●							
Funding acquisition			●				●		●				●	●	
Investigation	●	●		●	●	●		●		●	●	●			●
Methodology			●				●	●		●				●	
Project administration			●												
Software					●			●			●				
Resources	●		●				●		●	●			●	●	
Supervision			●				●	●	●				●	●	
Validation				●				●			●				
Visualization			●	●				●							
Writing – original draft			●					●							
Writing – review & editing			●				●	●	●				●	●	

Methods

Here, we first describe the processing steps in Suite3D. We then provide brief methods for the acquisition of the data that are analyzed in this study, and for the comparisons between analysis pipelines.

Suite3D

Suite3D is written in Python and is available at github.com/alihaydaroglu/suite3d/ under a GNU Affero GPL v3.0 open-source license (www.gnu.org/licenses/agpl-3.0.en.html). Below we describe its processing steps, and we provide pseudocode.

Preprocessing

Large-FOV 2-photon microscopes often use a resonant-galvo-galvo (RGG) scanning configuration for extended lateral scanning [5,25]. This results in the field of view being split into "strips", which often overlap at their edges. Suite3D uses the cross-correlation between edges of imaging strips to find the optimal shift between strips before fusing them together to construct a single continuous image.

Temporally multiplexed imaging methods such as Light Beads Microscopy can introduce inter-plane crosstalk since the inter-voxel time interval is similar to the fluorescence half-life of calcium indicators. We assume the following unidirectional model of crosstalk, where F_i is the true fluorescence and S_i is the measured signal in plane i, and α is a scalar "crosstalk coefficient":

$$S_i = F_i + \alpha F_j$$

For a user provided set of cross-talking pairs $\{(i_1, j_1), \dots (i_P, j_P)\}$, the default algorithm in Suite3D finds α that minimizes second derivative of the correlation coefficient with respect to the crosstalk coefficient, across all pairs:

$$\underset{\alpha}{\operatorname{argmin}} \sum_p \left| \frac{\partial^2}{\partial \alpha^2} \frac{\widehat{S_{i_p}}\, S_{j_p}}{||\widehat{S_{i_p}}\, S_{j_p}||} \right| \quad \text{where } \widehat{S_i} = S_i - \alpha S_j$$

We find that this approach outperforms minimizing the correlation coefficient directly, both in simulated and real data. This is due to similarities in structural features (blood vessels, neuropil) which create a non-zero correlation coefficient between the true signal in any pair of planes (F_i, F_j). Minimizing the correlation coefficient directly, by contrast, causes an overestimate of the crosstalk coefficient. Alternatively, the user can set a parameter to minimize the mutual information between S_i, S_j, similar to the PICASSO method [26].

The final step of the preprocessing phase is the construction of a "reference volume". Often, volumetric data is acquired as a series of axially separated planes between which there may be constant lateral offsets due to optical considerations. Suite3D uses a subset of the recording (typically 200-400 frames) to register planes to one another and undo inter-plane lateral offsets. Then, an iterative version of the volumetric motion correction algorithm (described in the next section) is applied to the subset of frames to construct a reference volume. This reference volume is often "sharper" than a simple time-averaged version of the movie because the brain motion has been corrected, so it is later used as a template during motion correction.

In practice, the computation of the preprocessing parameters is merged into a single "initialization" step which computes the optimal fusing and crosstalk values, constructs a reference volume and computes several quality metrics of the recording within several minutes on a subset of the full movie. In a typical workflow, this serves as a checkpoint where the user has the option to use the GUI to verify the recording

quality, exclude noisy planes and confirm processing parameters before moving onto the more time-consuming stages. In future processing steps, whenever a batch of frames is loaded from disk, the described preprocessing steps (fusing, crosstalk subtraction, plane-to-plane registration) are applied with precomputed parameters as part of the I/O pipeline.

3D motion correction

Rigid volumetric motion-correction is implemented with a phase-correlation based algorithm, similar to one previously described in Ref. [11]. We replaced the Fast Fourier Transform (FFT) step with a three-dimensional version and extended the multiplicative and additive spatially tapered masks to three dimensions to reduce artifacts at the edges of the volume. The resulting 3D phase correlations can be interpreted similarly to 2D phase correlation maps, where the location of the peak indicates the three-dimensional displacement that must be applied to the given frame. These displacements are applied as integer shifts of the entire volumetric frame. Optionally, the user can also apply a plane-by-plane non-rigid motion correction after the non-rigid step – this step computes a x and y shift for each tile of a plane, and uses these to compute a smooth field of shifts over the entire plane, and corrects them using bilinear interpolation [11]. We found that in higher noise, lower-resolution recordings, applying non-rigid shift can be detrimental to the data quality.

To accelerate motion correction, we developed a novel GPU-based implementation of the rigid and non-rigid motion correction algorithms. Previous attempts to accelerate the Suite2p motion correction algorithm with GPUs did not yield speed improvements, due to the costly nature of transferring large image data between RAM and GPU VRAM. To minimize the back-and-forth data transfers, we implement rigid (2D and 3D) and non-rigid (2D) registration steps directly on the GPU using a combination of CuPy-accelerated Python code and custom CUDA kernels in C++.

For a typical use case of Suite3D in a laboratory environment with active acquisition of large-scale volumetric data, one of the major bottlenecks is the file transfer speed. The raw data is typically moved from the experiment workstation to a shared server, from where it is typically moved to a fast disk on a processing workstation. Running pipelines such as Suite2p or CaImAn on data that is located on a server can cause significant slowdown, depending on I/O speeds. Suite3D motion correction implements a multi-threaded file I/O pipeline, where the next batch is loaded from storage into memory while the current batch is processed on the GPU, thus removing the requirement of moving data to local storage and masking the I/O times.

3D detection

The correlation volume is inspired by the 2D correlation map implemented in Suite2p. Before the motion-corrected movie is passed to the algorithm, it is temporally binned – the size of the temporal bin is a user parameter that should be set to the number of frames in a typical transient (*detection_timebin*). Next, the binned movie is temporally high-pass filtered with a rolling mean filter with a ~10s window to exclude low-temporal frequency variations. Next, each pixel is normalized by the smoothness of its temporal activity, estimated by the standard deviation of its first temporal differences, to account for varying signal-to-noise levels. To remove the neuropil-driven shared variability across pixels, the movie is spatially high-pass filtered with a cutoff around 70 μm to create the neuropil-subtracted movie. Next, the movie is spatially low-pass filtered with a cutoff around 5 μm to enhance structure at the functional length scales. These filters can be changed depending on the targets. For example, when imaging axonal boutons, the cutoff for the functional filter should be set to ~1-2 μm. On this filtered movie, the root-sum-square of each pixel across time is computed, only considering frames where the value of the given pixel was above the activity threshold, reducing the movie across the time dimension. The resulting correlation map is then thresholded and used as the seed for cell segmentation. Suite3D allows the user to sweep a range

of values for the sizes of the filters, the value of the threshold, and the strength of the normalization on a subset of the movie quickly, to manually optimize over the default settings if necessary for a particular dataset.

Suite3D implements several substantial algorithmic changes relative to Suite2p. First, the 2D spatial filters for neuropil subtraction and cell detection are extended to 3D. This allows improved detection for cells extending over several planes, and better estimation of neuropil at each voxel. The user has the option to select between Gaussian and top-hat filters for each. Second, Suite3D implements a local thresholding method [27] which adaptively thresholds the correlation map at each location. This approach is beneficial when imaging large field-of-views or using microscope configurations that may lead to uneven SNR at different imaging depths, where the correlation map will scale differently for each region of the imaged volume. This can be problematic for the segmentation step, as a single hard threshold across the full volume will be too low for some regions and too high for others. Third, Suite3D introduces an adjustable pixel normalization parameter that scales the strength of normalization. This makes the temporal-difference normalization more robust. Otherwise, in high-noise recordings, dark regions (such as blood vessels) can become bright in the correlation map. This artifact is remedied by setting the normalization parameter to 0.7-0.9.

The memory and processing requirements of computing the correlation map in one go as implemented in Suite2p (github.com/MouseLand/Suite2p) are often unfeasible for large volumetric recordings. Suite3D, therefore, computes the correlation map with a batch-wise algorithm. For running estimation of standard deviation across batches, we used Welford's algorithm [28]. The filtering and reduction steps are parallelized across CPU cores. To handle large arrays which cannot be held in memory at once, Suite3D uses efficient memory-mapped, parallelized algorithms supported by Dask (docs.dask.org/en/stable/) and Zarr (zarr.readthedocs.io/en/stable/) .

Algorithm 1: Batch-wise Correlation Map Computation		
Input:		
$\mathbf{M} \in \mathbf{R}^{t \times z \times y \times x}$		Motion corrected movie
Parameters		
t_f	`temporal_hpf`	Rolling temporal top-hat filter size (s)
$f_n \in \mathbf{R}^3$	`npil_filt_um`	Size of spatial neuropil filters in each dimension (μm)
$f_c \in \mathbf{R}^3$	`cell_filt_um`	Size of spatial cell detection filters (μm)
s	`sdnorm_exp`	Strength of pixel normalization (often 0.6-1.0)
i	`intensity_thresh`	Threshold for a pixel to be considered "active"
d	`detection_timebin`	Temporal binning on movie before computation
b	`batch_size`	Temporal batch size
Computation:		
$\mathbf{V}, \mathbf{A}, \mathbf{X}, \boldsymbol{S} \in \mathbf{R}^{z \times y \times x}$		Initialize correlation volume, mean and max volumes, and the normalizing volume
for each $\mathbf{B} = \mathbf{M}[t_i : t_i + b]$		Take a temporal batch of movie of size b
$\mathbf{A} = \frac{t_i}{t_i + b}\mathbf{A} + \frac{1}{b}\sum_{j=1}^{b}\mathbf{B}[j]$		Update the mean volume
$\mathbf{X} = \max(\max(\mathbf{B}), \mathbf{X})$		Update the maximum volume
$\mathbf{B} = \mathbf{B} - \mathbf{B} * \mathrm{rect}\left(\frac{t}{t_f}\right)$		Convolve movie along time axis with a rectangular filter, and subtract the result from the movie
$\boldsymbol{S} = \boldsymbol{S} + \frac{1}{b}\sum_{t}(\mathbf{B}[t_i] - \mathbf{B}[t_{i-1}])^2$		Update the normalization volume, which is the running sum of the variance of the first difference of each pixel

$\boldsymbol{\sigma} = \sqrt{\boldsymbol{S}}$

$\mathbf{B} = \mathbf{B} \,/\, \boldsymbol{\sigma}^{s}$

Normalize each pixel by the std. dev. of its temporal first differences, raised to exponent s which controls the strength of the normalization

$\mathbf{B}_{\mathrm{n}} = \mathbf{B} * g(f_n)$

Convolve the movie with a gaussian kernel to compute the low-pass filtered neuropil movie

$\mathbf{B}_{\mathrm{s}} = \mathbf{B} - \mathbf{B}_{\mathrm{n}}$

Subtract neuropil activity from original movie

$\mathbf{B}_{\mathrm{c}} = \mathbf{B}_{\mathbf{s}} * \ g(f_c)$

Low-pass filter with a cell-sized kernel to detect cell-sized structures

$\mathbf{V} \leftarrow \mathbf{V} + \frac{1}{b}\sum_{t} \mathbf{B}_{\mathrm{c}}[\mathbf{B}_{\mathrm{c}} > i]^{\mathbf{2}}$

For each pixel, compute the variance of its activity at timepoints where it is active to reduce the 4D movie into a 3D volume

end

$\mathbf{C} = \sqrt{\mathbf{V}}$

$\boldsymbol{C}_{\mathrm{thresh}} = \text{local thresholding } (\mathbf{C})$

3D segmentation

The segmentation phase iteratively selects candidate voxels from the correlation volume and expands them into ROI masks based on the correlation of their activity with neighbors. First, a seed voxel is selected from the correlation volume by taking the maximum and initializing a region of interest (ROI) of size 1. Next, the neighboring voxels are taken as candidate voxels for expanding this ROI. The candidate pixel selection and growth of the ROI happens in three dimensions, allowing Suite3D to compute volumetric cell masks. The activity of each candidate voxel is compared to the activity of the ROI in the neuropil subtracted movie saved in the detection step. The comparison uses only frames where the existing ROI is considered "active", by thresholding its activity by a hard threshold or 99$^{\text{th}}$ percentile cutoff, whichever is lower (both parameters are user-adjustable). Suite3D computes the correlation coefficient of each candidate voxel and the ROI and compares it to a set threshold to determine whether a candidate voxel is included. Each included voxel is weighted by their correlation with the ROI. We find that this algorithm is more robust to noise than the method implemented in Suite2p, and that it improves the segmentation results. Suite3D also provides the user an option to revert to the Suite2p-style method where the mean over active frames is used to determine expansion. This expansion step is executed iteratively (see Algorithm 2) until an ROI cannot be expanded anymore, after which a new seed voxel is selected for the next ROI.

The Suite2p algorithm can fail in lower-resolution datasets by detecting "junk" ROIs that are too small, too large, or too disjointed to be real cells. Suite3D re-casts the algorithm to create several interpretable parameters, which are exposed through a parameter selection UI and can be adjusted these can solve the problems. $c_{\min}$ controls the number of ROIs detected by changing the endpoint of the segmentation step – if obvious ROIs are being missed, this parameter should be lowered. e_{thr} controls how eager the algorithm is to expand a given ROI: if ROIs are too small, this should be lowered, if ROIs extend too wide, this should be increased. p_{active} and i_{thr} jointly determine the active frames of each ROI – for short recordings, lowering the thresholds tend to give better results, while for longer recordings (~hours) the thresholds can be set very high so only high-quality timepoints are considered as "active" frames.

Algorithm 2: Parallelized volumetric cell segmentation

Inputs:

$\mathbf{M_s}$: (t, z, y, x)	Normalized, neuropil subtracted movie ($\boldsymbol{B}_s$ from Algorithm 1, concatenated across all batches)
$\mathbf{C}$: (z, y, x)	Locally thresholded correlation map

Parameters:

c	`peak_thresh`	Minimum peak value of the correlation map
p	`percentile`	Percentile of activity above which a frame should be considered "active" for each ROI
a	`activity_thresh`	Threshold of z-scored activity above which a frame should be considered "active" for ROI
e	`extend_thresh`	Threshold of correlation above which a candidate pixel will be added to the ROI

Computation:

$\text{ROIs} \leftarrow [\,]$	
while $\max(\mathbf{C}) > c$:	
$s_z, s_y, s_x = \text{argmax}(\mathbf{C})$	Find the coordinates of the maximum value of the correlation map, and use it as the seed voxel
$s = \text{extend}(s_z, s_y, s_x)$	Initialize ROI as the seed voxel and its immediate neighbors
$w = \frac{\mathbf{1}}{\text{len}(s)^2}$	Initialize unit norm weight vector for each voxel in ROI
for $\boldsymbol{k}$: 1 to 3 do	
$m = \mathbf{M_s}[\boldsymbol{s}] \cdot w$	Compute weighted sum ROI voxels in movie
$v = \min(i_{\text{th}}, \text{pct}(m, p))$	Get threshold for "active" frames for this ROI, minimum of the hard threshold and the percentile threshold
$t = \text{where}(\mathbf{M_s}(s) > v)$	Identify active frames of ROI
while s is growing **do**	Extend ROI
$\hat{s} = \text{extend}(\mathbf{s})$	Get neighbors of ROI voxels as "candidate" voxels
$m_{\hat{s}} = \mathbf{M_s}[\hat{s}][\text{t}]$	Compute activity of candidate voxels on active frames
$e = \text{corr}(m_{\hat{s}}, m[t])$	Compute correlation coefficient of ROI and candidate voxels on active frames
$s = \hat{s}[\text{where}(e > e_{\text{thr}})]$	Update ROI voxels to include candidate voxels with correlation above threshold
$\text{update}(w)$	Weight each voxel by its covariance with the ROI
end while	
end for	
$\text{ROIs} = [\text{ROIs}, s]$	Add current ROI to list of ROIs after extending
$\mathbf{C}[s] = 0$	Remove ROI from correlation volume

Computationally, the segmentation algorithm is demanding as it requires the entire time series of the movie to be held in memory at once. To make it tractable, the movie is split into patches in the x and y dimensions, and each patch is loaded entirely into memory one at a time. The neuropil-subtracted movie used is the one computed in the correlation map computation step, which is located in a temporary file in fast local storage and accessed through memory-mapping. To account for cells that may cross boundaries between patches, small overlaps between patches are enforced and cells are de-duplicated at the end of the segmentation step. For a given spatial batch, the algorithm described is not straightforward to parallelize since the segmentation of each single requires updating the correlation map, which is then required to identify the next cell to segment. To parallelize across n CPU cores, our

algorithm greedily selects the top n peaks of the correlation map with a minimum distance constraint between all pairs of selected peaks. Then, each worker accesses the subtracted movie located in shared memory to segment, and each updates the correlation map in shared memory once their selected cell is extracted. This is implemented using the multiprocessing module, and specifically the SharedMemory class in Python.

Postprocessing and curation

Once weighted masks are computed for each region of interest, Suite3D returns to the motion-corrected movie to extract the weighted average of fluorescence over all pixels in each ROI. In addition, Suite3D implements a novel volumetric neuropil correction – a hollow shell around each ROI is computed (excluding other ROIs), and the average activity within this hollow shell is taken to be the neuropil fluorescence. This neuropil fluorescence is subtracted from the ROI fluorescence for each ROI with a coefficient of 0.7. Next, the neuropil-subtracted traces are deconvolved using a version of the OASIS algorithm with no L0/L1 constraints [19,20].

Once ROIs are extracted, a custom GUI based on napari [29]is used for semi-automated curation and quality control. The interactive GUI allows histogram-based filtering of ROIs based on size, values on the correlation volume, or activity levels, and is easily extensible to include user-defined cell filters. It also allows manual curation of individual ROIs, allowing the user to mark them as cells and non-cells. For high-quality 3D visualization, Suite3D exports cell maps in the .mrc format for visualization in Chimera [30], which was used for visualizations in Figure 5.

Datasets

To test and validate Suite3D, we acquired datasets at University College London, and we analyzed datasets previously obtained at Rockefeller University, University of California Berkeley, and University of Sussex.

University College London

Experiments at University College London (UCL) were conducted according to the UK Animals Scientific Procedures Act (1986) under personal and project licenses released by the Home Office following appropriate ethics review. The experiments were performed on 5 adult mice (3 female, 2 male; aged 11-25 weeks) expressing GCamP6s in all cortical excitatory neurons (CamKII x Ai162).

Mice were implanted with cranial windows of 4mm diameter under surgical anesthesia. Details of the surgery can be found in Ref. [31]. After recovery (minimum 5 days), mice were head-fixed and free to run on a low-resistance treadmill during imaging. In some recordings, visual stimulation was delivered via three screens covering approximately 270° horizontally and 70° vertically of the visual field of the animal. Drifting gratings were presented with 12 directions (30° spacing) in random order, with eight repeats of each unique stimulus. Each grating was presented for 2 s, with a 2-3 s inter-stimulus interval.

Light Bead Microscopy (LBM) recordings at UCL were conducted with a modified Many-fold Axial Multiplexing Module (MAxiMuM) cavity [7], integrated with a commercial mesoscope (Thorlabs, Multiphoton Mesoscope). The laser source comprised of a 4.8 MHz pump laser (Coherent Monaco, 60W average power at 1030 nm) with an optical parametric chirped-pulse amplifier (OPCPA, White Dwarf from Class 5 Photonics), with a maximum average output power of 6W at 960 nm.

Conventional multi-plane imaging was conducted on a commercial 2-photon microscope (Bergamo II, Thorlabs). Sequential acquisition of planes was achieved by piezo-driven actuation of the objective. Visual cortex imaging was conducted on transgenic mice expressing GCaMP6s in Layer 5 excitatory

neurons (Rbp4 x Ai162), with similar cranial windows as described above. Hippocampal imaging was conducted on mice with viral expression of GCaMP6f in CA1, imaged through an optical cannula.

Rockefeller University

LBM recordings at Rockefeller University were conducted in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC) on adult transgenic mice (2 female; aged 10-20 weeks) expressing GCamP6s in excitatory neurons (vGlut1-Cre x Ai162). The imaging was performed on a similar LBM as described above, using similar methods.

University of California Berkeley

FACED microscopy at University of California Berkeley was conducted in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC). The dataset that was analyzed was obtained in a mouse expressing GCamP6s in excitatory neurons (Slc17a7-IRES2-Cre × Ai162D). The imaging was done with a lateral voxel spacing of 1.4 µm and 2 µm in x and y, and an axial spacing of 5 µm. Further details can be found at [4,6].

University of Sussex

High-resolution imaging of retinal boutons at the University of Sussex was conducted according to the UK Animals Scientific Procedures Act (1986) under personal and project licenses released by the Home Office following appropriate ethics review. The dataset that was analyzed was obtained in a mouse expressing SyGCaMP8m virally [32]. A conventional 2-photon microscope was used, with lateral voxel spacing of 0.5 µm and an axial spacing of 2 µm on mice. Further details can be found in [23].

Comparison of pipelines

To evaluate the performance of Suite3D, we compared it plane-by-plane Suite2p[11,12] and to volumetric CaImAn [13].To apply plane-per-plane Suite2p and volumetric CaImAn directly on LBM data, the preprocessing steps described earlier were applied to fuse imaging strips together, subtract crosstalk and save the data in a format usable by either algorithm. For Suite2p we used version 0.14.4, as implemented in github.com/MouseLand/Suite2p, and we ran it with default parameters independently on each plane. For CaImAn we used version 1.11.4, available at github.com/flatironinstitute/CaImAn. To run it, we used the workflow in the tutorial notebook for volumetric imaging in the CaImAn documentation, adapting parameters relating to neuron size and timescale.

To match cells detected in Suite2p and Suite3D, we used a semi-automated method considering the distance between the centroid of each pair of cells across modalities, and the correlation of their extracted fluorescence. To ensure that matches were genuine, we chose conservative thresholds for correlation and distance. Then, we estimated fraction of shot noise [21] for the extracted fluorescence traces for matched ROIs across cells.

To compare variability across trials, we normalized each fluorescence trace by its standard deviation and extracted responses in the time windows where stimuli were presented. We took the standard deviation of each ROI's fluorescence across eight repeats of the same stimulus, averaged over all unique stimuli.

To compute the number of significantly stimulus-responsive cells in each pipeline, we considered all ROIs (matched and non-matched). For each ROI, we took the average activity in the time period from 0.5s after stimulus onset to 2.0s after stimulus onset. We computed the fraction of signal-related variance across two halves of repeats of each stimulus as described in [31]. To identify statistically significant cells, we repeated the same computation 500 times, shuffling the stimulus identities. Cells with a signal-

related variance fraction above the 99th percentile of the null distribution were considered significantly responsive.

All computation was conducted on a Linux workstation with an Intel Xeon w9-3475 CPU with 36 dual cores (72 hyper-threading cores) operating at a base frequency of 2.2 GHz, with 512 GB of DDR5 RAM, and an NVIDIA RTX A4500 GPU with 20 GB of VRAM. The workstation was equipped with a pair of 4TB M.2 SSDs in RAID0 configuration and had a 10 Gbps connection to the server on which raw data was stored. Parallel processing options were enabled for all pipelines tested. For Suite2p and CaImAn, the time required to load movies from the server, preprocess them into a suitable format and save them to the fast local disk before processing is included under "file I/O" in Figure 4g.

References

1. Göbel, W. & Helmchen, F. In Vivo Calcium Imaging of Neural Network Function. *Physiology* **22**, 358–365 (2007).

2. Rupprecht, P., Prendergast, A., Wyart, C. & Friedrich, R. W. Remote z-scanning with a macroscopic voice coil motor for fast 3D multiphoton laser scanning microscopy. *Biomed. Opt. Express* **7**, 1656–1671 (2016).

3. Stringer, C. *et al.* Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, 255 (2019).

4. Lai, Q. T. K. *et al.* High-speed laser-scanning biological microscopy using FACED. *Nat. Protoc.* **16**, 4227–4264 (2021).

5. Wu, J. *et al.* Kilohertz two-photon fluorescence microscopy imaging of neural activity in vivo. *Nat. Methods* **17**, 287–290 (2020).

6. Zhong, J. *et al.* FACED 2.0 enables large-scale voltage and calcium imaging in vivo. *Manuscsript Rev.* (2025).

7. Demas, J. *et al.* High-speed, cortex-wide volumetric recording of neuroactivity at cellular resolution using light beads microscopy. *Nat. Methods* 1–9 (2021) doi:10.1038/s41592-021-01239-8.

8. Bellafard, A., Namvar, G., Kao, J. C., Vaziri, A. & Golshani, P. Volatile working memory representations crystallize with practice. *Nature* **629**, 1109–1117 (2024).

9. Bouchard, M. B. *et al.* Swept confocally-aligned planar excitation (SCAPE) microscopy for high-speed volumetric imaging of behaving organisms. *Nat. Photonics* **9**, 113–119 (2015).

10. Voleti, V. *et al.* Real-time volumetric microscopy of in vivo dynamics and large-scale samples with SCAPE 2.0. *Nat. Methods* **16**, 1054–1062 (2019).

11. Pachitariu, M. *et al.* Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv* 061507 (2017) doi:10.1101/061507.

12. Stringer, C., Pachitariu, M. & Contributors. suite2p. (2025).

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

14. Dinc, F. *et al.* Fast, scalable, and statistically robust cell extraction from large-scale neural calcium imaging datasets. Preprint at https://doi.org/10.1101/2021.03.24.436279 (2021).

15. Manley, J. *et al*. Simultaneous, cortex-wide dynamics of up to 1 million neurons reveal unbounded scaling of dimensionality with neuron number. *Neuron* (2024) doi:10.1016/j.neuron.2024.02.011.

16. Paramo, B., Bachmann, S. O., Baudouin, S. J., Martinez-Garay, I. & Davies, A. M. Neuregulin-4 Is Required for Maintaining Soma Size of Pyramidal Neurons in the Motor Cortex. *eneuro* **8**, ENEURO.0288-20.2021 (2021).

17. Yan, C. *et al*. Automated and Accurate Detection of Soma Location and Surface Morphology in Large-Scale 3D Neuron Images. *PLOS ONE* **8**, e62579 (2013).

18. Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nat. Methods* **2**, 932–940 (2005).

19. Friedrich, J., Zhou, P. & Paninski, L. Fast online deconvolution of calcium imaging data. *PLOS Comput. Biol.* **13**, e1005423 (2017).

20. Pachitariu, M., Stringer, C. & Harris, K. D. Robustness of Spike Deconvolution for Neuronal Calcium Imaging. *J. Neurosci.* **38**, 7976–7985 (2018).

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

22. Kerlin, A. M., Andermann, M. L., Berezovskii, V. K. & Reid, R. C. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. **Neuron**, (2010).

23. Baruchin, L. J., Alleman, M. & Schröder, S. Reward Modulates Visual Responses in the Superficial Superior Colliculus of Mice. *J. Neurosci.* **43**, 8663–8680 (2023).

24. Rossi, L. F., Harris, K. D. & Carandini, M. Spatial connectivity matches direction selectivity in visual cortex. *Nature* **588**, 648–652 (2020).

25. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *eLife* **5**, e14472 (2016).

26. Seo, J. *et al*. PICASSO allows ultra-multiplexed fluorescence imaging of spatially overlapping proteins without reference spectra measurements. *Nat. Commun.* **13**, 2475 (2022).

27. Gonzalez, R. C. *Digital Image Processing Using MATLAB / Rafael C. Gonzalez, Richard E, Woods, Steven L. Eddins*. (McGraw Hill Education India, New Delhi, 2010).

28. Welford, B. P. Note on a Method for Calculating Corrected Sums of Squares and Products. *Technometrics* **4**, 419–420 (1962).

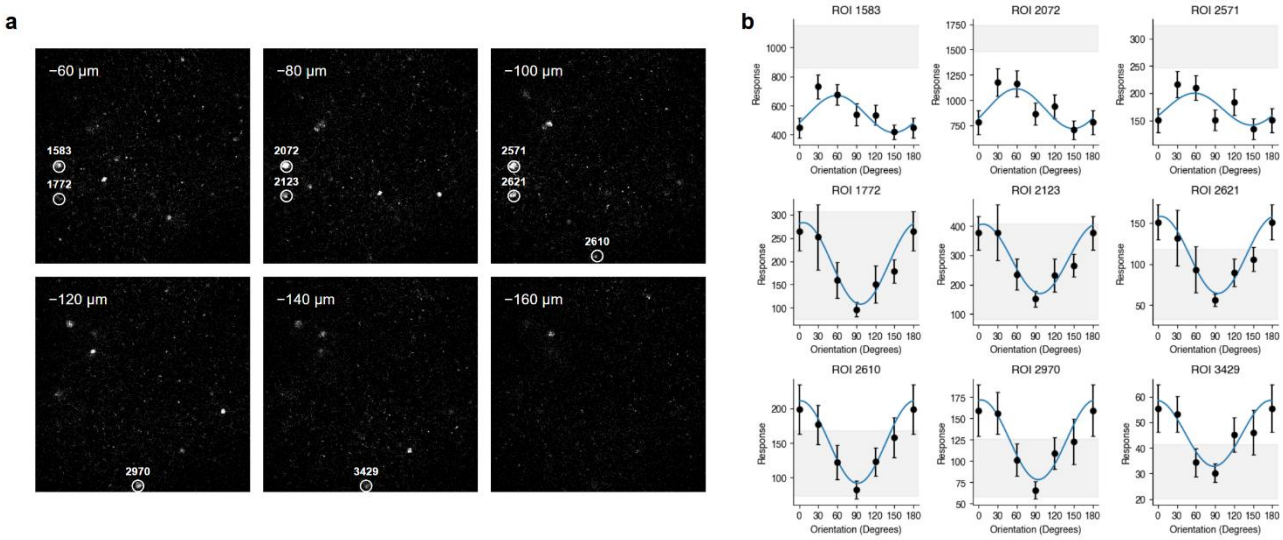
29. Sofroniew, N. *et al*. napari: a multi-dimensional image viewer for Python. (2025) doi:10.5281/zenodo.15029515.

30. Pettersen, E. F. *et al*. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

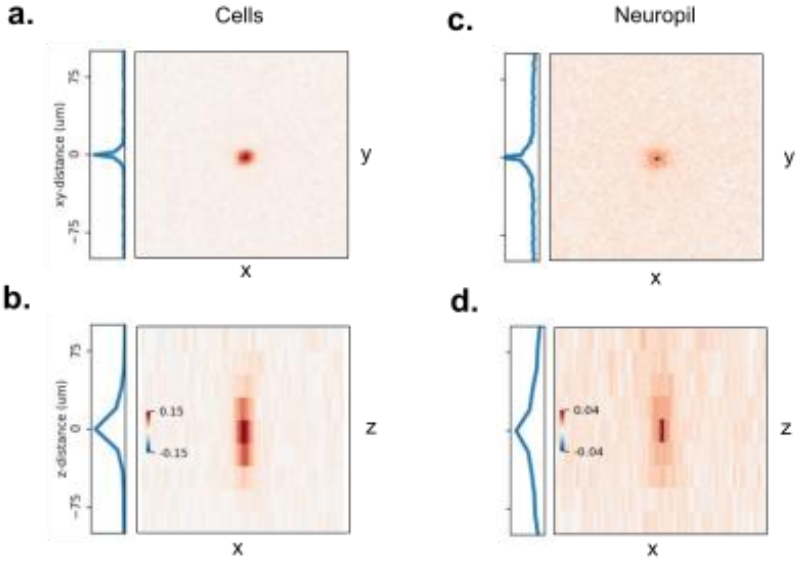
31. Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M. & Harris, K. D. High-dimensional geometry of population responses in visual cortex. *Nature* **571**, 361–365 (2019).

32. Zhang, Y. *et al*. Fast and sensitive GCaMP calcium indicators for imaging neural populations. *Nature* **615**, 884–891 (2023).

Supplementary Figures



Supplementary Figure 1. **Cortical pyramidal cells can extend over 40 µm axially when imaged with a conventional 2-photon microscope**. We used a multi-plane 2-photon microscope (Thorlabs, Bergamo-II) to image neurons in the visual cortex of mice expressing GcaMP6s in all excitatory neurons and tdTomato in a sparse set of those neurons. Planes were 20 µm apart**.** We then used Suite2p to analyze the GcaMP6 images plane by plane. **a.** Images obtained in red from structural tdTomato marker in 6 planes. Labels indicate select ROIs detected by Suite2p in the functional movie, overlaid on the sparse structural image. The 9 circled ROIs are likely to be the same 3 neurons appearing across planes. **b.** Orientation tuning curves for ROIs labelled in (a). ROIs appearing in the same planar position have identical tuning curves, indicating that they are the same neuron, split across multiple planes and overcounted with plane-by-plane analysis.



Supplementary Figure 2. **Volumetric data contains 3D correlation structures. a,** The average fluorescence correlation of a voxel in a cell with its neighbors in x and y for an LBM recording. **b,** Same as (a), shown in x and z directions. **c,** Average fluorescence correlation of a neuropil voxel (e.g. a voxel not located within a cell) with its neighbors in x and y for the same recording as (a). **d**, Same as (c), in x and z.