Suite2p: beyond 10,000 neurons with standard two-photon microscopy

Marius Pachitariu[1,2*], Carsen Stringer[3], Sylvia Schröder[4], Mario Dipoppa[1,2], L. Federico Rossi[4], Matteo Carandini[4] and Kenneth D. Harris[1,2]
1 — UCL Institute of Neurology, London WC1E 6DE, United Kingdom.
2 — UCL Department of Neuroscience, Physiology, and Pharmacology, London WC1E 6DE, United Kingdom.
3 — Gatsby Computational Neuroscience Unit, London W1T 4JG, United Kingdom.
4 — UCL Institute of Ophthalmology, London EC1V 9EL, United Kingdom.
* — Correspondence to marius.pachitariu.10@ucl.ac.uk

Abstract

The combination of two-photon microscopy recordings and powerful calcium-dependent fluorescent sensors enables simultaneous recording of unprecedentedly large populations of neurons. While these sensors have matured over several generations of development, computational methods to process their fluorescence remain inefficient and the results hard to interpret. Here, we introduce a set of practical methods based on novel clustering algorithms, and provide a complete pipeline from raw image data to neuronal calcium traces to inferred spike times. We formulate a generative model of the fluorescence image, incorporating spike times and a spatially smooth neuropil signal, and solve the inference and learning problems using a fast algorithm. This implementation scales linearly with the number of recorded cells, and the complete pipeline runs in approximately one hour for typical two-hour long recordings, on commodity GPUs. Furthermore, this method recovers twice as many cells as a previous standard method. This allowed us to routinely record and detect $\sim$10,000 cells simultaneously from the visual cortex of awake mice using standard two-photon resonant-scanning microscopes. The software is publicly available at github.com/cortex-lab/Suite2P, together with a graphical user interface that allows rapid manual curation of the results.

1 Introduction

Standard, commercial two-photon microscopes readily image the activity of large numbers of neurons, but algorithms for processing the resulting data still suffer from significant limitations. Ideally, such algorithms should satisfy several criteria. First, they should be fast, to keep up with ever-larger data sets produced by next-generation microscopes[1,2]. Second, their output should be accurate, so that a human operator need only spend little time curating the algorithm's result. Third, they should generalize to recordings of multiple cell types and brain regions, which can exhibit widely different activity patterns. Fourth, they should model and appropriately handle experimental confounds such as neuropil contamination[3]. Finally, it would be ideal if the algorithms could run on inexpensive workstations rather than require a cluster of servers, as some current software packages do[4].

Here we demonstrate a set of fast and accurate algorithms which fulfil these criteria, providing a complete pipeline from raw data (Fig 1a-e) to activity traces and spikes. The output of the pipeline is a set of interpretable deconvolved signals that represent single spikes or bursts. The

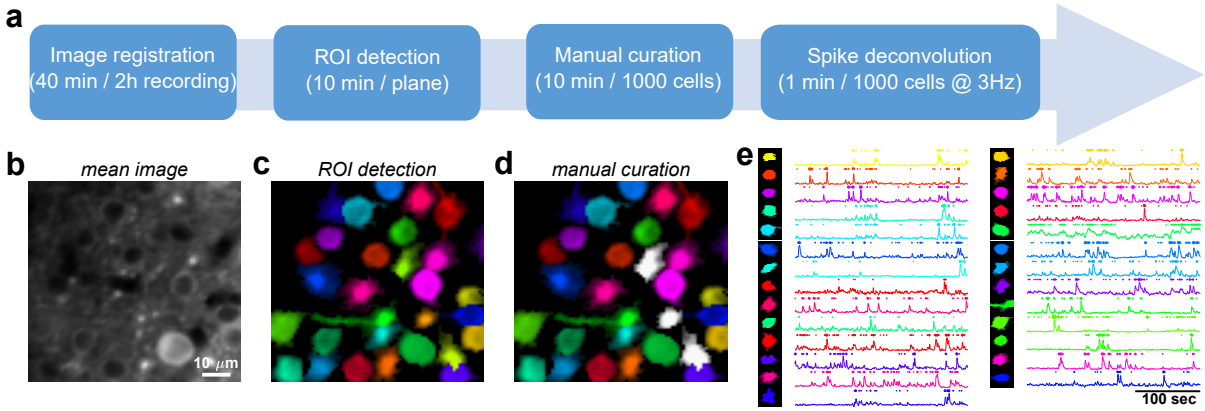

Figure 1. Pipeline schematic. a, Pipeline schematic containing four major steps: image registration, cell detection, manual curation and spike deconvolution. The processing time required for each is indicated. **b,** Mean image of a segment of the field of view. **c,** Cluster assignments of the algorithm are indicated by color (chosen randomly for each cluster). **d,** Manual selection and removal of non-somatic ROIs (see full GUI below). **e,** Fluorescence traces of the cells shown in (d), together with their deconvolved spike times, represented as dots of area proportional to the deconvolved amplitude of the burst.

modelling framework explicitly includes the effects of neuropil contamination. Indeed, we show that the neuropil contributes a large amount of variance to the overall recorded signals, but that the statistics of its contribution allow it to be efficiently and correctly removed from spiking traces. Finally, the entire pipeline runs in approximately real-time, on a single workstation equipped only with a consumer-grade GPU*.

The software is available at https://github.com/cortex-lab/Suite2P, and it includes a graphical user interface (GUI) that allows a human operator to quickly curate the results of the automated pipeline and discard misclassified signals. It thus provides a fast and reliable end-to-end system for two-photon data processing.

To illustrate the performance of this system we demonstrate how we used it to obtain >10,000 cells from imaging data obtained with a standard resonant-scanning two-photon microscope. A video example of the raw data is available at www.youtube.com/watch?v=xr-flH2Ow2Y, and the results are shown in Figure 2, representing the cells detected by the algorithm. To validate the results, we test the performance of the automated pipeline using simulated data with the same statistics as our recordings, and on ground truth data from simultaneous cell-attached recordings.

1.1 Previous work

Most two-photon data processing methods work by grouping together pixels with correlated time-courses. Such methods can in general be formalized as a problem of matrix decomposition, combined with sparsity or dynamical penalties.

*For the analyses in this paper we used a workstation with an Intel i7 processor, 32 GB of RAM, and a GTX 970 or 980Ti GPU

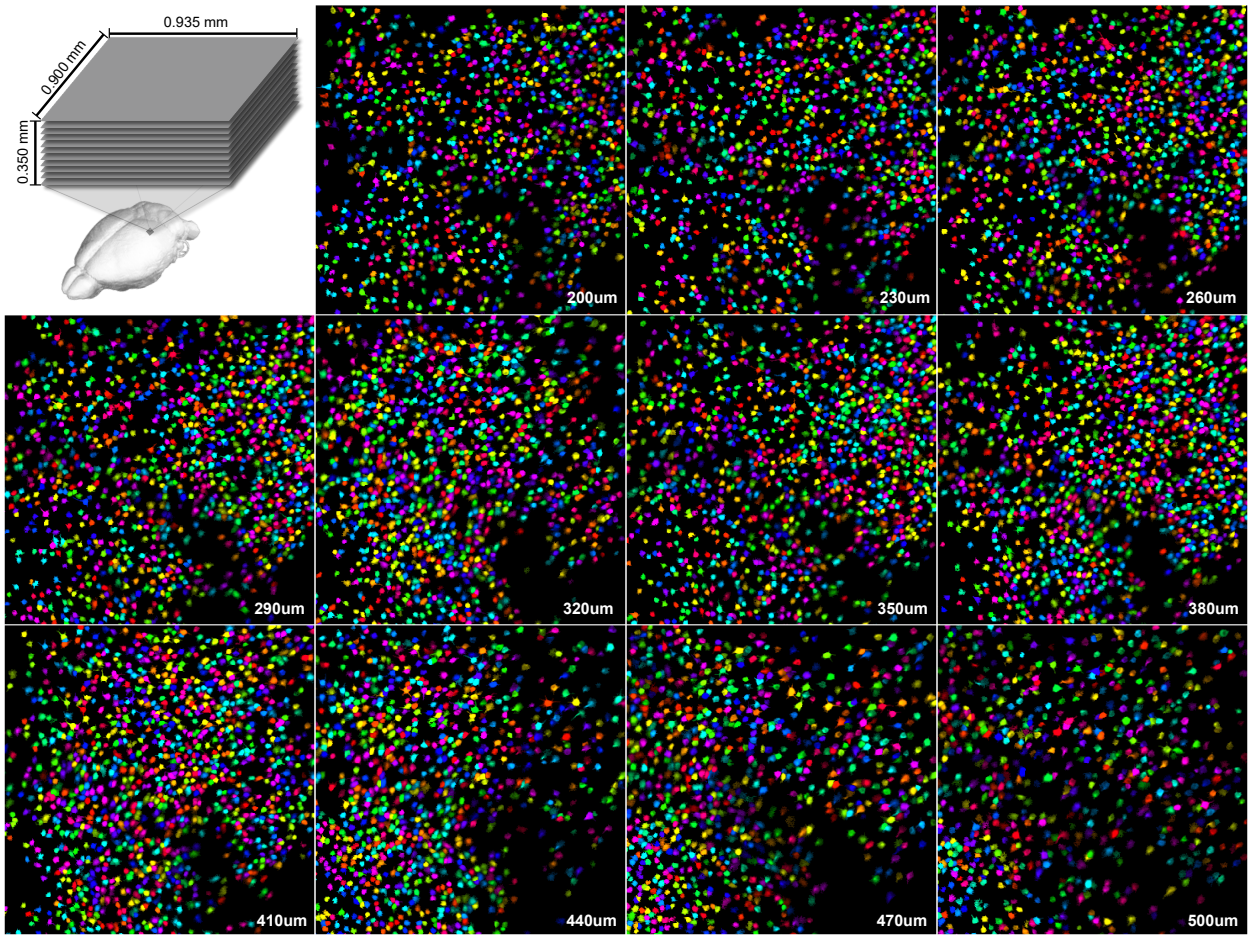

Figure 2. 13,451 simultaneously recorded neurons. We recorded 11 planes in the visual cortex of a transgenic mouse (Ai94, Camk2a-tTA, EMX Cre triple tg; GCaMP6s in pyramidall cells in all layers) at 2.5Hz, using a standard Thorlabs B-scope with resonant scanning. The pseudocolor masks of all detected cells are shown here. A short segment of the raw fluorescence in all planes is shown at https://www.youtube.com/watch?v=xr-flH2Ow2Y.

One of the earliest automated approaches used standard algorithms for independent components analysis (ICA), which demixes the "sources" of a multi-dimensional signal in order to maximize a cost function, in this case skewness[5]. For small fields of view, the algorithm can be run quickly by applying it to a low-dimensional PCA projection of the data. This method, however, becomes impractical for typical fields of view in cortex which commonly involve hundreds to thousands of regions of interest (ROIs). In this case the field of view needs to be broken down into sections and the algorithm run on each separately, extending the run time and requiring arbitration for overlapping patches. Moreover, the ICA approach does not explicitly model the neuropil signal, other than forcing its estimate to be decorrelated from the estimated neural activity. However, as we illustrate below, the neuropil component will often be highly correlated with individual cells' activity: assuming decorrelation thus leads to an incorrect separation of neuropil from cellular signals.

Another matrix decomposition approach finds components based on positivity constraints on the

cell masks combined with a model of the calcium dynamics[6]. This model captures the activity of a cell as an auto-regressive process with one or two decay timescales, to be fit from data. The cost function is cast into a convex optimization problem, which guarantees global maxima and allows the pipeline to be implemented with standard optimization algorithms. However, solving a generic convex optimization problem is slow on large data sets such as those we consider here. Furthermore, as we illustrate below, these methods do not necessarily model the pervasive neuropil signal in a correct fashion.

An alternative approach to factorization-based methods consists of greedily segmenting neighboring pixels of high correlation[7–9]. Such methods may be a practical choice for very large datasets, as they can run much faster than most existing factorization-based approaches. Nevertheless, it remains open whether such methods can perform as well as factorization-based methods. Our implementation of such an approach (called autoROI) is compared to the new method below.

Finally, some alternative methods[10], instead find cell locations based on an anatomical image, an average taken over all frames of the movie. Although fast, this approach does not provide precise cell boundaries and often misses sparsely active cells with low baseline fluorescence. In systems such as the cerebral cortex, such cells constitute the rule rather than the exception.

2 Model formulation

Our algorithm is based on a generative model of the two-photon movies, derived from simple notions of the underlying biological structure and of the optics of the microscope. The model expresses the recorded fluorescence movie in terms of two underlying signals: a set of spike trains, each associated with a discrete ROI (the cellular soma or other compartment); and a neuropil signal that is assumed to vary slowly across the field of view. Each pixel's fluorescence trace is a sum of the neuropil signal at that location, plus a contribution from ROIs.

2.1 Pre-processing: registration via phase correlations

Before the model is fit, a first preprocessing step is performed, to spatially register all frames in the movie to each other. The most common registration technique used in two-photon microscopy relies on finding the cross-correlation peak between a frame and a target image[11], which can be computed efficiently with fast Fourier transforms. We improved this method by spatially whitening the frame and the target prior to computing their cross-correlation. The resulting method is called phase correlation[12,13]. We tested it on simulated data with known translation, and we found that it substantially outperforms cross-correlation (not shown).

We also provide a non-rigid version of the method, by smoothly interpolating the XY offsets computed on sub-blocks of the image. This allows us to use the same efficient Fourier methods for batch processing, and simultaneously constrains the inferred flow map of the image to be smooth. We did not rely on more advanced optic flow methods, or affine transformation methods, due to efficiency concerns. The non-rigid method was applied successfully in experiments with rotational movements of up ~ 5 degrees (within plane and particularly outside plane), and for aligning fields of view recorded on different days for chronic imaging of the same set of cells.

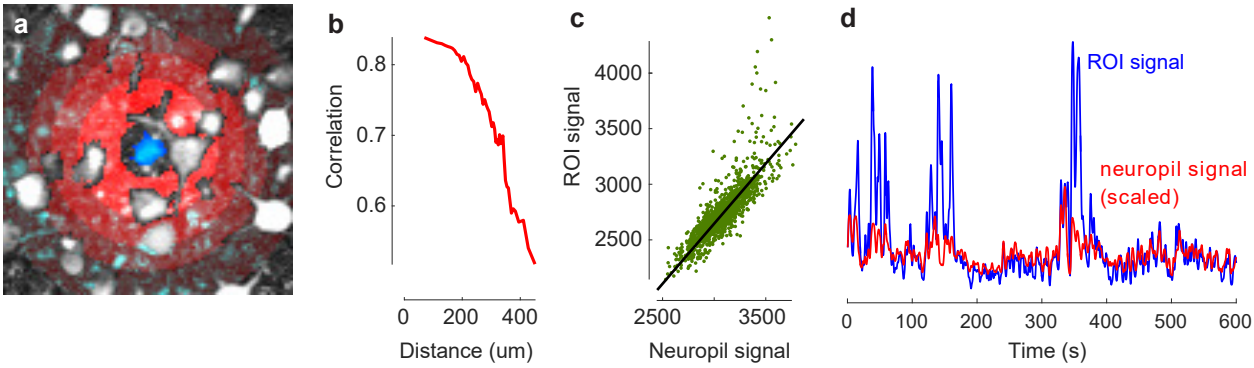

Figure 3. The statistics of neuropil contamination. a, Selected cell (blue) and neuropil areas (red-shaded annuli) superimposed on the variance image (white). The neuropil signal here is defined as the signal inside the red-shaded annuli but not in any detected ROIs. **b,** Correlation between the cell's fluorescence signal and the neuropil signal is large, and decays slowly with distance. **c** Scatter plot of cell and neuropil signals (green points), together with a robust regression fit (black line), that ignores outliers (above fit line, which correspond to transients). **d,** Example timecourse of ROI signal and neuropil signal. The neuropil signal has been scaled by 0.93, as found by our algorithm.

In our pipeline, registration accounts for a large fraction of the running time, but can be sped up 4-fold by use of consumer-grade GPUs.

2.2 Modelling the neuropil contamination

Two-photon microscopes acquire signals with a point-spread function that is typically elongated by an order of magnitude more in the Z-axis than in X and Y. Thus, the signal recorded in a single plane is in fact a weighted average of a volume extending across tens of microns in Z, and although a cell might appear to be the only element occupying a spatial position, the signals recorded there are contaminated by signals from the extracellular space above or below it[3]. This volume will contain sometimes another cell, but most often the neurites (axons and dendrites) of very many other cells, adding a contamination signal that represents the averaged activity of a large cell population.

To account for this "neuropil" signal in the model, we must first understand its spatial distribution. By analyzing pixels outside of ROIs, we found that the neuropil signal is very highly correlated in space, and can be assumed to be correlated over regions of several tens of microns (Fig. 3a-b, data from visual cortex). Furthermore, the neuropil signals were clearly highly correlated with signals recorded over the somas, and the transients of the cells were seen as distinct deviations from an otherwise linear relationship between a somatic signal and the surrounding neuropil (Fig. 3c; note also that these deviations occurred primarily when the neuropil signal was itself large, reflecting increasing probability of cellular firing at times of high network activity). This emphasizes that the spatially smooth neuropil signal is indeed still present at the soma and accounts for a considerable fraction of the recorded signal[3,14]. Some methods[6] model the neuropil as a one-dimensional signal shared by all pixels with different weights, but others[5] do not enforce any constraints. We represent the neuropil in a set of spatially-localized basis functions (raised cosines), thus allowing the neuropil signal to vary slowly across space (see below).

Although the timecourse of the neuropil varies smoothly in space, the scaling of this signal in

different pixels varies. In particular, a cell displaces some of the neuropil mass, thus the amount of neuropil signal inside a cell is typically less than the amount of contamination. In turn, cells that are more out of focus will also contain a larger fraction of neuropil signal. How should we determine the true level of contamination and correctly demix the signal at the soma from the signal just outside the soma? The distinguishing feature of cellular activity is its relatively sparse timecourse convolved with an asymmetric temporal kernel. Thus, by using a model for the cell's activity that incorporates this knowledge, we can find the subtraction coefficient which makes the signal most likely to originate from a single cell.

2.3 Generative model

The model for the recorded signal r_k at pixel k is

$$\begin{aligned} r_k &= \sum_n \Lambda_{kn} \cdot \mathbf{f}_n + \alpha_k \sum_i B_{ki}\mathbf{n}_i + \eta \\ \Lambda_{kn} &> 0, \text{if pixel k belongs to ROI}_n \\ \Lambda_{kn} &= 0, \text{otherwise} \\ \eta &= \mathcal{N}(0, \sigma). \end{aligned} \tag{1}$$

Here Λ describes the (very sparse) matrix of pixel membership, $\mathbf{f}_n$ is the fluorescence timecourse for ROI$_n$, $\sum_i B_{ki}\mathbf{n}_i$ is a model of the neuropil at pixel k, represented in a fixed set of spatial basis functions B, $\mathbf{n}_i$ is the timecourse of each neuropil component, and α_k denotes the fraction of the neuropil signal that contaminates pixel k. If no large somatic compartment fills the point-spread function then α_k will be close to 1. The basis functions B do not need to be fit and are fixed as a set of isotropic 2d raised cosine functions. The raised cosines tile the full FOV, with inter-center spacing of ten times the diameter of a cell.

3 Cell detection optimization

Before the model can be used to find cells and estimate the underlying spike trains, several parameters need to be inferred from the data. We use a fitting strategy loosely based on the standard expectation-maximization (EM) algorithm. The algorithm alternates between an E step, where it estimates the spatial distribution of the biological sources contributing to the measured fluorescence (cell ROIs and neuropil), and an M step where it estimates the temporal dynamics of these sources and the underlying spike trains.

The cost function which needs to be minimized is

$$\text{Cost}(\mathbf{n}, \mathbf{s}, \mathbf{b}, \mathbf{k}, \alpha, \Lambda, B) = \left\| r_k - \sum_n \Lambda_{kn}\mathbf{f}_n - \alpha_k \sum_i B_{ki}\mathbf{n}_i \right\|^2, \tag{2}$$

such that $\Lambda_{kn} \geq 0$. Furthermore, we typically constrain the temporal kernel $\mathbf{k}$ to be equal for all cells in a given recording, which is a good approximation and prevents overfitting individual kernels to very sparsely firing cells. However, we also allow these kernels to be fitted independently if, for example, different cell classes are imaged together with potentially different Calcium dynamics.

At the E-step, we assume we are given a subset of the parameters in equation 2, and attempt to fit the other ones. Specifically, we are given $\mathbf{n}, \mathbf{s}, \mathbf{b}, \mathbf{k}$, which represent the underlying cellular and neuropil activity, and we optimize $\mathbf{\Lambda}$ and α, which together define the cluster memberships and the scaling factors for ROI and neuropil signals respectively. If a given pixel can only belong to one ROI, such as in a pure clustering problem, then we can easily find its cluster membership as the index $\delta(k)$ that minimizes the quadratic form 2 with two free linear parameters $\Lambda_{k\delta(k)}$ and α_k. If a pixel is allowed to belong to multiple ROIs, then consecutive cluster assignments are found by matching pursuit[15]. For simplicity, none of the results shown here consider more than one cluster assignment per pixel.

At the M-step, we assume we are given $\mathbf{\Lambda}$ and α, and determine the continuous-time ROI signals $\mathbf{f}$ (defined in equation 1) and the continuous-time neuropil traces $\mathbf{n}$ by linear regression. Note that the fluorescence $\mathbf{f}$ is not constrained to be positive, or to fit the generative model, except in the final iteration of the algorithm, see next section.

We alternate the E and M step several dozen times before the procedure converges. After convergence, all detected ROIs are split into their spatially-connected components. We classify as putative somatic ROIs those that lie within a certain user-specified range (e.g. 1/3 to 3 times as many pixels as an "average cell"), and simultaneously have their pixels relatively close the centroid of the ROI (relative to the most compact shape: a disk). To avoid local minima, we typically anneal the number of clusters during optimization by a factor of 2, using pre-specified values, though the precise values do not matter much for the end results because most extra components are absorbed into dendritic clusters.

3.1 Accuracy of cell detection

The new pipeline was able to detect a large number of cells in our two-photon data: compared to a previous system (based on detection of locally correlated pixels[7,8]), it detected up to two times more cells (Fig. 4).

To determine the accuracy of the new method in detecting cells, we created a simulated ground truth benchmark. We first ran the pipeline on a single experiment and manually curated the results. We then chose a small subset of cells (20 out of 300) and added their activity to a neighboring spatial location of the movie, which did not overlap significantly with the rest of the ROIs (Fig. 5a-b). We then re-ran the pipeline and compared the results with the 20 known cell locations. We re-ran this test 43 times, choosing a different set of cells every time. The analysis shows that almost all ground truth ROIs were correctly identified (Fig. 5c). A small proportion of ROIs only had partial matches of their spatial masks, which represented cells that were over-split by the algorithm, but whose activity timecourse was nonetheless identified correctly (Fig. 5c). Note that over-split ROIs can be easily merged manually by a user, while over-merged ROIs are much harder to split. Our results suggest that very few ROIs were over-merged, since such merged cells would get very poor scores by our metric (Fig. 5c).

Finally, we analyzed what factors influenced detection accuracy (number of pixels, cell compactness, signal variance and signal skewness, Fig. 5d). We correlated these factors with the detection score for each ROI. Cells with large number of pixels were typically broken into multiple compo-

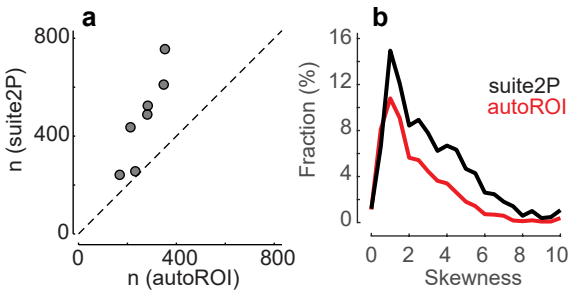

Figure 4. Suite2P detects more cells. a. Total cell yield comparison on 5 datasets processed with AutoROI (a method based on[7,8]) and Suite2P. **b.** Histogram of cell yields as a function of the skewness of the fluorescence signal.

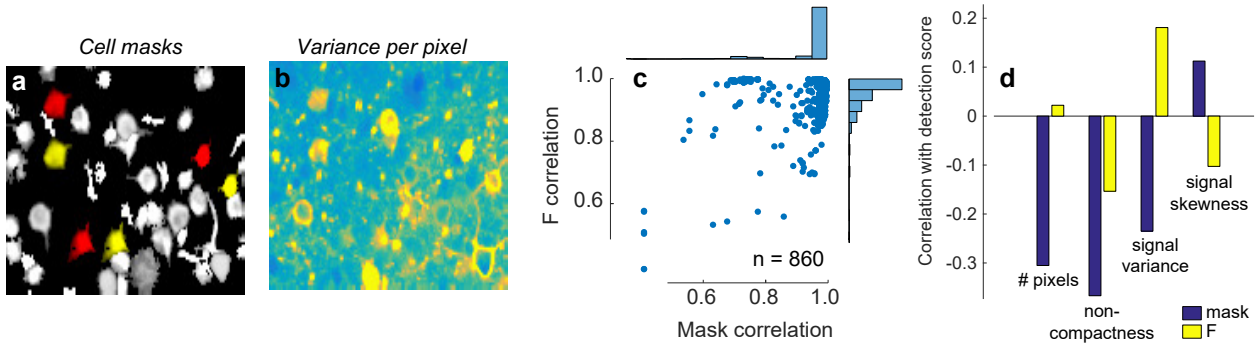

Figure 5. Cell detection performance. a. Illustration of hybrid ground truth generation. A small percentage of cells, detected by a human operator, were added to the raw movie in the neighborhood of their original location. Yellow cells are the source masks and red cells are the new location where the ground truth cells were added. **b,** Variance map of the new movie. **c,** Scatterplot of detection accuracy measured as mask shape accuracy and activity timecourse accuracy. **d,** Factors influencing cell detection performance.

nents and scored lower in terms of spatial mask recovery. However their recovered fluorescence signals do not appear to have suffered, likely because each sub-component faithfully followed the timecourse of the original cell. Similarly, cells that were significantly less compact than a disk were less well recovered (independently of total number of pixels). Signal variance and skewness played less of a role in the quality of cell detection.

4 Spike deconvolution

For spike deconvolution, we use L0 priors to define the contribution of a burst of spikes to the calcium trace, unlike the L1 priors[6,16] previously used. Thus our deconvolved spike trains are very sparse, with 90 to 99% of timepoints set to 0. Solving the optimization problem for L0 constraints is typically a hard problem, but is made simpler here by the structure of the neural activity: calcium transients are sparse and identifiable by their asymmetric shape. Consequently, a simple matching pursuit algorithm can identify them.

We denote by $\mathbf{s}_m(t)$ the true underlying spiking process of ROI m, which for convenience is discretized at the time-resolution of the two-photon movie and thus represents the total number of spikes in a timebin. We model the fluorescence trace $\mathbf{f}_m$ of an ROI as arising from the convolution of $\mathbf{s}$ with a temporal kernel $\mathbf{k}$:

$$\mathbf{f}_m = \mathbf{s}_m * \mathbf{k} + \beta_m \mathbf{p}_m + b_n \tag{3}$$

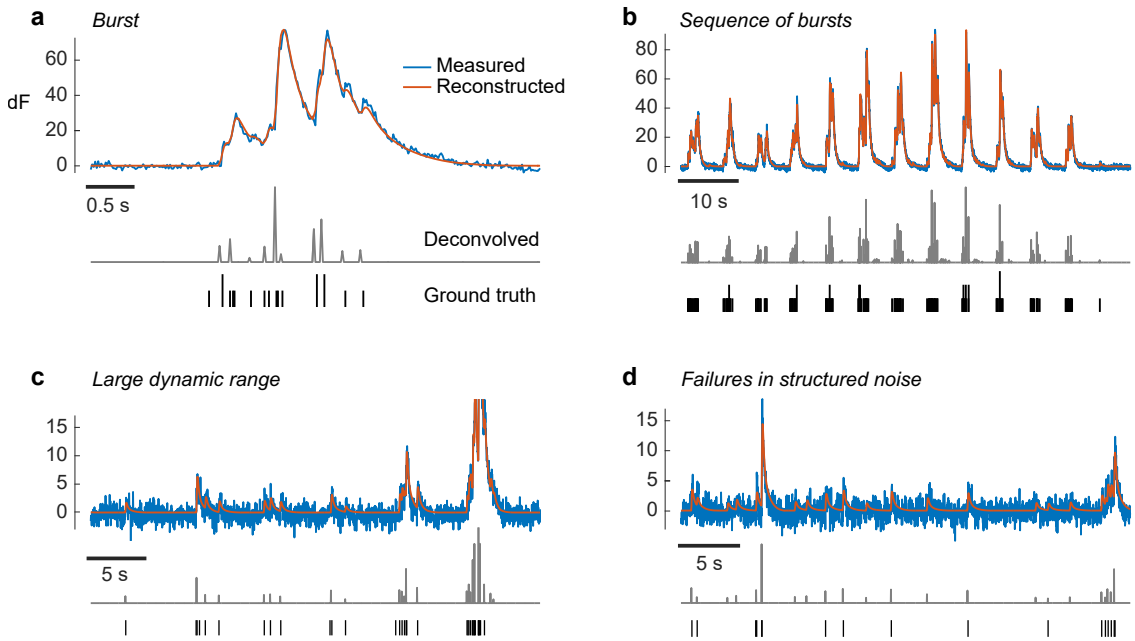

Figure 6. Example spike deconvolution of ground truth data. a, Example trace of a neuron recorded simultaneously with two-photon calcium imaging and with a cell-attached pipette. Ground truth spikes were binned at the time resolution of the calcium imaging data, and displayed as black ticks of length proportional to the number of spikes per bin. **b,** Zoom out of **a,** to show more bursts. **c,** Example consecutive periods with very different spiking rates. **d,** Example segment with a lower signal-to-noise ratio. In this case the algorithm sometimes detected spikes in the noise.

where b_n is a baseline parameter representing resting fluorescence in the ROI[3], $*$ denotes temporal convolution and $\mathbf{p}_m$ is the timecourse of the neuropil contamination assigned to this cell by the cell detection model. For regularization and smoothness purposes, we represent the kernel $\mathbf{k}$ as a weighted sum of exponentially-decaying basis functions, with timescales chosen to be consecutive powers of 2.

We determine the spike trains $\mathbf{s}$, temporal kernel $\mathbf{k}$, and baseline fluorescence values $\mathbf{b}$, using a custom matching pursuit algorithm, which sequentially introduces new spikes into the approximation, while iteratively refining the times of the already selected spikes, until a threshold is reached (the threshold is defined relative to the variance of the signal). We also re-estimate the neuropil contamination coefficient β_m, to take advantage of the increased modelling capacity of the full dynamical model of spiking-dependent Calcium.

4.1 Accuracy of spike deconvolution

We evaluated the spike deconvolution method on ground truth data from simultaneous two-photon and electrical recordings[3] (Fig. 6). Our results on this dataset are comparable to those of a recent state-of-the-art method[17], though we did not perform a direct comparison. Unlike previous methods[6,16,17], our L0 deconvolution returns traces with extreme sparseness values, with non-zero values narrowly distributed around the times of the true spikes (Fig. 7).

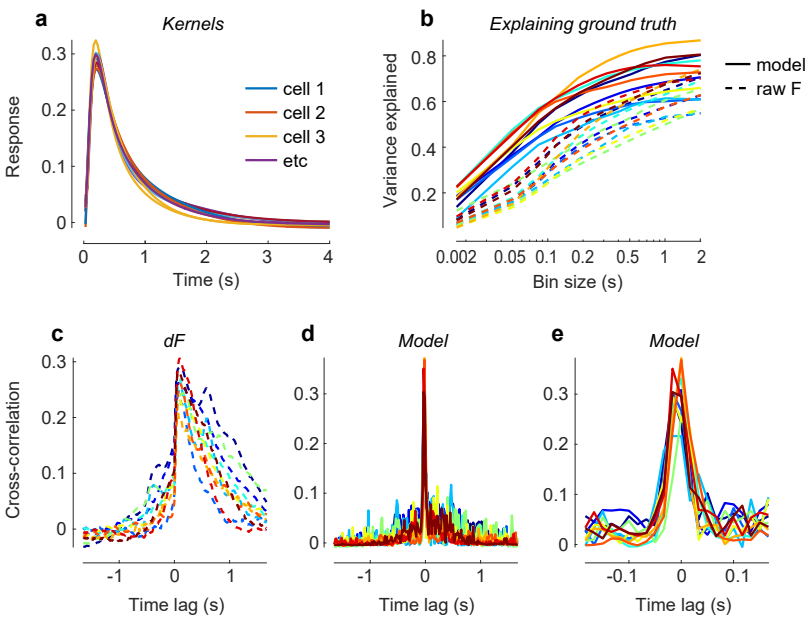

Figure 7. Quantification of spike deconvolution results. **a,** Temporal kernels recovered by the model on 10 cells expressing the indicator GCaMP6f (four of which are also shown in Fig. 5). **b,** Variance of the binned ground truth spike data explained by the model (solid) and raw fluorescence (dashed). The best performance for raw fluorescence was obtained by (further) temporal smoothing and finding the best temporal delay. **c,** Cross-correlogram between ground truth spikes and recorded fluorescence (dF). **d,** Cross-correlogram between ground truth spikes and model-deconvolved spikes. **e,** Same as (d), but zoomed in 10 times to show the temporal precision of the deconvolution.

5 Accelerating the algorithm by large-scale SVD decomposition

The execution speed of the algorithm is greatly boosted by a simple computational trick. After spatial registration, we reduce the dimensionality of the data using an SVD decomposition, and perform all EM iterations in a space of much lower dimension than the original movies. This step greatly reduces the space required to store the data, denoises the data, and allows our modelling framework to operate fast enough for large-scale recordings (10-100x depending on original number of frames in the dataset).

We approximate the recorded fluorescence $\mathbf{F}$ (represented as a matrix with one row for each pixel and one column for each timepoint) of all pixels k at times t with its best approximation $\mathbf{F} = \mathbf{U}\mathbf{V}^T$, where $\mathbf{U}$ has as many rows as pixels, $\mathbf{V}$ has as many rows as timepoints, and both have a limited number of columns. The singular values have been absorbed into $\mathbf{U}$, which is orthogonal but not orthonormal, while $\mathbf{V}$ is an orthonormal matrix.

The SVD reduction can be applied consistently in our model because the cost function is of the form

$$\begin{aligned}\text{Cost} &= (F - f_{\text{model}})^T (F - f_{\text{model}}) \\ &\approx (\mathbf{U}\mathbf{V}^T - f_{\text{model}})^T (\mathbf{U}\mathbf{V}^T - f_{\text{model}})\end{aligned}$$

Under the low-rank approximation of F, the model trace f_{model} can only have power in the V subspace, because subspaces orthogonal to V can only add noise to the cost function. Hence $f_{\text{model}} = u_{\text{model}} V^T$, and the cost function reduces to

$$\begin{aligned}\text{Cost} &\approx (U - u_{\text{model}})^T V^T V (U - u_{\text{model}}) \\ &\approx (U - u_{\text{model}})^T (U - u_{\text{model}}),\end{aligned}$$

Hence, as long as the cost function is quadratic, the SVD decomposition can be used to effectively reduce the dimensionality of the problem to a reasonable size. To be able to perform an SVD approximation on the type of large-scale data we acquire, we binned the signal at each pixel temporally until we reduced the number of timepoints to under 10,000, then computed the eigenvectors of the time by time covariance matrix, which were then used to determine the spatial singular vectors.

Once cell masks and neuropil weightings have been estimated using this SVD-accelerated approach, we recompute the fluorescence timecourses of each cell using the full data for additional accuracy.

6 Manual GUI

To curate the output of the algorithm, we implemented an intuitive GUI that shows large numbers of cells at the same time, thus allowing the human visual system to do a parallel search for detecting candidates to be labelled as cells or non-cells. An image of the GUI is shown in Fig. 8.

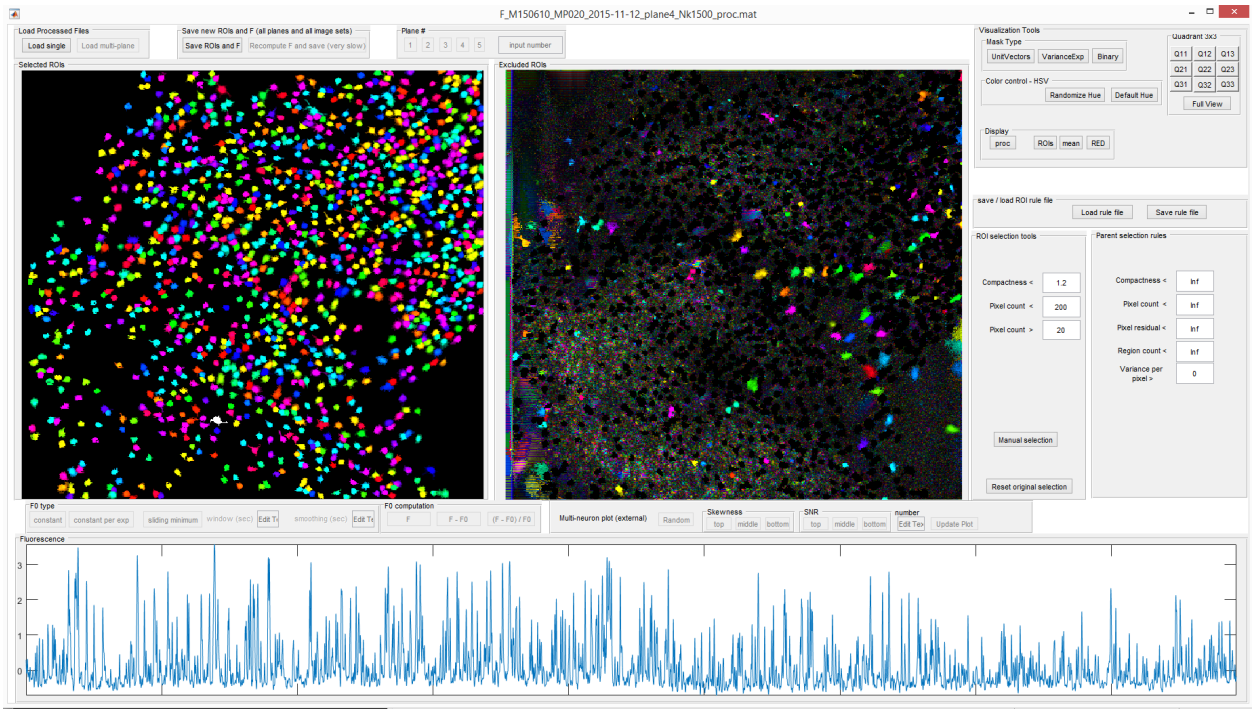

Figure 8. Graphical user interface. The main pseudocolor panels show ROIs identified by Suite2P. In this case, 1,276 active cells were identified from a single plane, in an 11 plane recording with a total yield of 13,451 active cells. The left panel shows the "accepted" ROIs, while the right panel shows the "rejected" ROIs. The GUI allows for curation of these automated results. Several options are provided for globally setting anatomical criteria for accepting ROIs. A point-and-click interface is used for overwriting the automated assignments. Detailed instructions for using the GUI are included online.

Figure 9. Responses to visual stimuli in 7167 simultaneously-recorded neurons. We recorded nine planes from layer 2/3 of mouse visual cortex in a transgenic GCaMP6s-expressing line (Camk2a-tTA; Ai94(TITL-GCaMP6s); Rasgrf Cre Triple Tg). We sorted all neurons by their preferred stimulus out of 112 natural images. The activity of all neurons is shown for a duration of approx 4 minutes, displayed as a raster plot of the deconvolved responses. Colored dots are shown at the times of the stimuli, just above the population of neurons which best-responded on average to that stimulus.

7 Discussion

We provide a complete set of tools for two photon processing, with provably good performance and very fast run times. The pipeline is modular, allowing the main processing steps to be used independently: registration, cell detection, spike deconvolution and manual GUI. We have successfully used the tools presented here to scale up the yields of two-photon recordings and automate most of the processing of $\sim$10,000 neurons recorded simultaneously at 3Hz sampling rates (Fig. 2). This has now become a routine experiment in our lab. Existing processing pipelines[5,6,16,17] would take days or weeks of computation[†] to transform these data into single-neuron activity traces or spikes. Although a case has been made for parallelizing such pipelines over large compute clusters[4,18], the use of such computing resources is not available to many neuroscience labs, except via expensive commercial rental.

Finally, we demonstrate robust and selective neural responses to stimuli on a trial-by-trial basis in our $\sim$7,000 neuron recordings, by displaying the data in a manner similar to electrophysiology

[†]We extrapolated these estimates from runs on small FOVs, which together tile the full FOV.

("raster" plots, Fig. 9). Each row represents a cell and the dots represent deconvolved spike times, scaled by the amplitude of the spike burst, as identified by the algorithm. Groups of cells responding to their preferred stimulus can be clearly seen as narrow vertical stripes, thus illustrating the temporal precision of the spike deconvolution method. Using multiplane recording protocols such as ours, in combination with efficient automated processing methods, will allow the characterization of complete neural populations, on a trial-by-trial basis.

Acknowledgements

We thank Charu Reddy for surgeries and Michael Krumin with assistance with the two-photon microscopes.

This work was supported by the Wellcome Trust (95668, 95669, 108726), and the Simons Foundation (325512). CS was funded by a four-year Gatsby Foundation PhD studentship. MD and SS were supported by Marie Skłodowska-Curie Fellowships. LFR was funded by a four-year Wellcome Trust PhD studentship in Neuroscience. MC holds the GlaxoSmithKline / Fight for Sight Chair in Visual Neuroscience

References

1. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *bioRxiv* 055947 (2016).

2. Stirman, J., Smith, I., Kudenov, M. & Smith, S. Wide field-of-view, multi-region two-photon imaging of neuronal activity in vivo. In *Optics and the Brain*, BTu2D–2 (Optical Society of America, 2016).

3. Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).

4. Freeman, J., Vladimirov, N., Kawashima, T., Mu, Y., Sofroniew, N. J., Bennett, D. V., Rosen, J., Yang, C.-T., Looger, L. L. & Ahrens, M. B. Mapping brain activity at scale with cluster computing. *Nature methods* **11**, 941–950 (2014).

5. Mukamel, E., Nimmerjahn, A. & Schnitzer, M. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* **63**, 747–760 (2009).

6. Pnevmatikakis, E. A., Soudry, D., Gao, Y., Machado, T. A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W. *et al.* Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* (2016).

7. Smith, S. L. & Häusser, M. Parallel processing of visual space by neighboring neurons in mouse visual cortex. *Nature neuroscience* **13**, 1144–1149 (2010).

8. Vladimirov, N., Mu, Y., Kawashima, T., Bennett, D. V., Yang, C.-T., Looger, L. L., Keller, P. J., Freeman, J. & Ahrens, M. B. Light-sheet functional imaging in fictively behaving zebrafish. *Nature methods* (2014).

9. Kaifosh, P., Zaremba, J. D., Danielson, N. B. & Losonczy, A. Sima: Python software for analysis of dynamic fluorescence imaging data. *Frontiers in neuroinformatics* **8** (2014).

10. Pachitariu, M., Packer, A. M., Pettit, N., Dalgleish, H., Hausser, M. & Sahani, M. Extracting regions of interest from biological images with convolutional sparse block coding. In *Advances in Neural Information Processing Systems*, 1745–1753 (2013).

11. Poort, J., Khan, A. G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., Bauza, M., Sahani, M., Keller, G. B., Mrsic-Flogel, T. D. *et al.* Learning enhances sensory and multiple non-sensory representations in primary visual cortex. *Neuron* **86**, 1478–1490 (2015).

12. Alba, A., Vigueras-Gomez, J. F., Arce-Santana, E. R. & Aguilar-Ponce, R. M. Phase correlation with sub-pixel accuracy: a comparative study in 1d and 2d. *Computer Vision and Image Understanding* **137**, 76–87 (2015).

13. Foroosh, H., Zerubia, J. B. & Berthod, M. Extension of phase correlation to subpixel registration. *Image Processing, IEEE Transactions on* **11**, 188–200 (2002).

14. Peron, S. P., Freeman, J., Iyer, V., Guo, C. & Svoboda, K. A cellular resolution map of barrel cortex activity during tactile behavior. *Neuron* **86**, 783–799 (2015).

15. Mallat, S. & Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing* **41**, 3397–3415 (1993).

16. Vogelstein, J. T., Packer, A. M., Machado, T. A., Sippy, T., Babadi, B., Yuste, R. & Paninski, L. Fast nonnegative deconvolution for spike train inference from population calcium imaging. *Journal of neurophysiology* **104**, 3691–3704 (2010).

17. Theis, L., Berens, P., Froudarakis, E., Reimer, J., Roman-Roson, M., Baden, T., Euler, T., Tolias, A. S. & Bethge, M. Benchmarking spike rate inference in population calcium imaging. *Neuron* **90**, 471–482 (2016).

18. Freeman, J. Open source tools for large-scale neuroscience. *Current opinion in neurobiology* **32**, 156–163 (2015).