eNeuro

Research Article: Methods/New Tools | Novel Tools and Methods

Automatic Cell Segmentation by Adaptive Thresholding (ACSAT) for large scale calcium imaging datasets

Simon P. Shen[1], Hua-an Tseng[2], Kyle R. Hansen[2], Ruofan Wu[3], Howard Gritton[2], Jennie Si[3] and Xue Han[2]

[1]*Department of Physics, Harvard University, Cambridge, MA 02138, USA*
[2]*Biomedical Engineering Department, Boston University, Boston, MA 02215, USA*
[3]*School of Electrical, Computer and Energy Engineering, Tempe, AZ 85287, USA*

DOI: 10.1523/ENEURO.0056-18.2018

Received: 3 February 2018

Revised: 16 August 2018

Accepted: 23 August 2018

Published: 4 September 2018

Author Contributions: S.P.S. designed research, performed research, contributed unpublished reagents/ analytic tools, analyzed data, and wrote the paper; H.T. performed research, analyzed data, and wrote the paper; K.R.H. performed research; R.W. analyzed data; H.G. performed research; J.S. and X.H. designed research and wrote the paper.

Funding: NIH Director's new innovator award
1DP2NS082126

Funding: Pew Foundation

Funding: DARPA Young Faculty Award

Funding: Boston University Biomedical Engineering Department

Funding: National Science Foundation Graduate Research Fellowship
DGE-1247312

Conflict of Interest: Authors report no conflict of interest.

X.H. acknowledges funding from NIH Director's new innovator award (1DP2NS082126), Pew Foundation, DARPA Young Faculty Award, Boston University Biomedical Engineering Department. K.R.H. is supported by a National Science Foundation Graduate Research Fellowship under Grant No. DGE-1247312.

S.P.S. and H.T. are co-first authors

In the "Materials and Methods: Mouse preparation" section, "All animal procedures were approved by [Author University = Boston University] Institutional Animal Care and Use Committee."

In the "Materials and Methods: Code accessibility" section, "The code/software described in the paper is freely available online at [redacted for double-blind review = www.github.com/sshen8/acsat]."

Correspondence address: Jennie Si, Tel.: 480-965-6133; E-mail: si@asu.edu; Xue Han, Tel.: 617-358-6189; E-mail: xuehan@bu.edu

Cite as: eNeuro 2018; 10.1523/ENEURO.0056-18.2018

Automatic Cell Segmentation by Adaptive Thresholding (ACSAT) for large scale calcium imaging datasets

Running Title: Auto Cell Segmentation by Adaptive Thresholding

Simon P. Shen[1]*, Hua-an Tseng[2]*, Kyle R. Hansen[2], Ruofan Wu[3], Howard Gritton[2], Jennie Si[3**], Xue Han[2**]
[1] Department of Physics, Harvard University, Cambridge, MA 02138
[2] Biomedical Engineering Department, Boston University, Boston, MA 02215
[3] School of Electrical, Computer and Energy Engineering, Tempe, AZ 85287
*S.P.S. and H.T. are co-first authors

Contributions:
S.P.S. designed research, performed research, contributed unpublished reagents/analytic tools, analyzed data, and wrote the paper; H.T. performed research, analyzed data, and wrote the paper; K.R.H. performed research; R.W. analyzed data; H.G. performed research; J.S. and X.H. designed research and wrote the paper.

** X.H. and J.S. are co-correspondence
J.S.: 480-965-6133; si@asu.edu
X.H.: 617-358-6189; xuehan@bu.edu

Number of Figures: 11
Number of Tables: 0
Number of Multimedia: 2
Number of words for Abstract: 209
Number of words for Significance Statement: 118
Number of words for Introduction: 749
Number of words for Discussion: 1713

Authors report no conflict of interest

X.H. acknowledges funding from NIH Director's new innovator award (1DP2NS082126), Pew Foundation, DARPA Young Faculty Award, Boston University Biomedical Engineering Department. K.R.H. is supported by a National Science Foundation Graduate Research Fellowship under Grant No. DGE-1247312.

In the "Materials and Methods: Mouse preparation" section, "All animal procedures were approved by [Author University = Boston University] Institutional Animal Care and Use Committee."

In the "Materials and Methods: Code accessibility" section, "The code/software described in the paper is freely available online at [redacted for double-blind review = www.github.com/sshen8/acsat]."

Automatic Cell Segmentation by Adaptive Thresholding (ACSAT) for large scale calcium imaging datasets

Abstract

Advances in calcium imaging have made it possible to record from an increasingly larger number of neurons simultaneously. Neuroscientists can now routinely image hundreds to thousands of individual neurons. An emerging technical challenge that parallels the advancement in imaging a large number of individual neurons is the processing of correspondingly larger datasets. One important step is the identification of individual neurons. Traditional methods rely mainly on manual or semi-manual inspection, which cannot be scaled for processing large datasets. To address this challenge, we focused on developing an automated segmentation method, which we refer to as Automated Cell Segmentation by Adaptive Thresholding (ACSAT). ACSAT works with a time-collapsed image and includes an iterative procedure that automatically calculates global and local threshold values during successive iterations based on the distribution of image pixel intensities. Thus, the algorithm is capable of handling variations in morphological details and in fluorescence intensities in different calcium imaging datasets. Within this manuscript we demonstrate the utility of ACSAT by testing it on 500 simulated datasets, two wide-field hippocampus datasets, a wide-field striatum dataset, a wide-field cell culture dataset, and a two-photon hippocampus dataset. For the simulated datasets with truth, ACSAT achieved over 80% recall and precision when the signal-to-noise ratio was no less than ~24 dB.

Significance

ACSAT aims at automatically segmenting cells in large-scale calcium imaging datasets. It is based on adaptive thresholding at both global and local levels and iteratively

identifies individual neurons in a time-collapsed image. It is designed to address a variety of datasets, potentially involving variations in cell morphology and fluorescence intensity between different datasets. We demonstrate the effectiveness of ACSAT by testing it under a variety of conditions. For the simulated datasets with truth, ACSAT achieved recall and precision rates over 80% when the signal-to-noise ratio was no less than ~24 dB. For the datasets from mouse hippocampus and striatum, ACSAT captured ~80% of human-identified ROIs and even detected some low-intensity neurons that were initially undetected by human referees.

Introduction

The ability to record from a large population of single neurons during behavior greatly facilitates the investigation of the contribution of individual neurons to neuronal network dynamics. Extracellular single-unit recording has traditionally been a method of choice in neurophysiological analyses of single neurons in the brain. Recent improvements, such as the new generation of genetically-encoded calcium sensors GCaMP6 (Chen et al., 2013, Sun et al., 2013), have made it possible to observe hundreds to thousands of individual neurons simultaneously (Ohki et al., 2005, Andermann et al., 2010, Huber et al., 2012, Ziv et al., 2013, Issa et al., 2014, Mohammed et al., 2016). Though indirect, these calcium indicators have been sensitive enough to monitor neuronal activity with high spatiotemporal precision in behaving animals, allowing researchers to examine the activity of populations of a specific cell type (Hofer et al., 2011, Wachowiak et al., 2013, Pinto and Dan, 2015, Allen et al., 2017) or the same cell over an extended period of time (Poort et al., 2015).

As the performance of genetically-encoded calcium indicators improved, wide-field microscopy has become feasible for recording the activity of a large population of neurons over an extended anatomical area (Lutcke et al., 2013, Wilt et al., 2013). Although lacking the spatial subcellular resolution of a multiphoton microscope, wide-field microscopes can operate at a higher speed, allowing the simultaneous recording of increasingly larger populations (Ghosh et al., 2011, Ziv et al., 2013, Kim et al., 2016, Mohammed et al., 2016). Advanced microfabrication techniques further miniaturized the wide-field microscope to a microendoscope capable of monitoring neural activity in freely-moving animals (Ghosh et al., 2011, Ziv et al., 2013).

An emerging technical challenge that parallels advances in calcium imaging is the processing of large datasets (Hamel et al., 2015). During data analysis, an important step is to identify regions of interest (ROIs) corresponding to individual neurons. As data grows rapidly both spatially and temporally, the traditional labor-intensive approach of manual inspection has to be automated. Principal component analysis (PCA) and independent component analysis (ICA) methods are natural and frequently-used candidates for automating ROI identification (Mukamel et al., 2009). However, if its assumption of statistical independence between neurons is violated, which is often the case in real neural recordings, then the method relies on user selection of parameters for spatial segmentation.

Threshold-based methods represent a promising and intuitive alternative for automatic ROI identification. However, several challenges need to be overcome, including variability in recording conditions or fluorescence signal strength across structures,

recording subjects, and the imaging field. For example, one of the most referenced thresholding methods, Otsu's method, which automatically selects the optimal threshold value that minimizes the intra-class variance among ROI pixels and among background pixels, would only successfully segment some of the highest-intensity ROIs (Otsu, 1979, Sezgin and Sankur, 2004). Additionally, the multi-class Otsu's method is limited because uneven lighting may result in separate background classes. A waterfall-thresholding approach addresses uneven lighting by iterative thresholding to capture all intensity peaks, but its selection of a threshold value is "*ad hoc*," making it dataset-dependent and user-dependent (Mellen and Tuong, 2009). A feedback loop-based approach for segmenting bacteria cells optimizes the threshold value from the distribution of pixel intensities, but its assumption that the total ROI area remains constant over time does not hold for calcium-imaging datasets because neurons change in brightness (Shen et al., 2015). A recent machine learning-based algorithm uses image gradients and pixel traces to optimize the threshold value, but it still requires a user's subjective input in selecting a background removal factor based on each dataset (Fantuzzo et al., 2017). Other approaches based on edge detection have trouble due to weak fluorescence signal strength in comparison with the background pixels (Sadeghian et al., 2009). Generally, most segmentation methods require a high level of tuning to each individual dataset.

To overcome these challenges of diverse imaging datasets, we introduce a new Automated Cell Segmentation by Adaptive Thresholding (ACSAT) algorithm. ACSAT dynamically and automatically determines global and local threshold values based on the distribution of pixel intensities within a time-collapsed image of a recorded image

sequence. We demonstrate the utility of ACSAT on simulated datasets, cell culture datasets, and in vivo wide-field and two-photon datasets. For the simulated datasets with truth, ACSAT achieved recall and precision >80% when the signal-to-noise ratio was no less than ~24 dB. ACSAT also captured ~80% of human-identified ROIs in datasets from mouse hippocampus and striatum and was even able to detect low-intensity neurons that were initially undetected by human referees.

Materials and Methods

Wide-field hippocampus and striatum datasets

All animal procedures were approved by [Author University – redacted for double-blind review] Institutional Animal Care and Use Committee. Female C57BL/6 mice (8-12 weeks old, Taconic, Hudson, NY) were first injected with 250nL AAV9-Syn-GCaMP6.WPRE.SV40 virus (titer: ~6e12 GC/ml, University of Pennsylvania Vector Core). AAV was delivered either into the dorsal CA1 (AP: -2, ML: 1.4, DV: -1.6), or into the dorsal striatum (AP: 0.5, ML: 1.8, DV: -1.6) regions. Injections were performed with a 10 µL syringe (World Precision Instruments, Sarasota, FL) coupled with a 33 gauge needle (NF33BL, World Precision Instruments, Sarasota, FL) at a speed of 40 nL/min, controlled by a microsyringe pump (UltraMicroPump 3-4, World Precision Instruments, Sarasota, FL). Upon complete recovery, a custom imaging chamber with glass coverslip was surgically implanted on top of the viral injection site by removing the overlying cortical tissue. The imaging chamber was assembled by fitting a circular coverslip (size 0; OD: 3 mm) to a stainless steel cannula (OD: 0.317 mm, ID: 0.236 mm) using a UV-curable optical adhesive (Norland Products). During surgery, a custom aluminum

153 headplate was also attached to the skull, which allows head-fixation during the imaging
154 session.

155 Imaging data were acquired with a custom wide-field microscope coupled with a
156 scientific CMOS camera (ORCA-Flash 4.0, C11440-42U, Hamamatsu, Boston MA),
157 controlled by the commercial software package HCImageLive (Hamamatsu, Boston
158 MA). The wide-field microscope consisted of a Leica N Plan 10×0.25 PH1 objective
159 lens, an excitation filter (HQ 470/50), a dichroic mirror (FF506-Di02), an emission filter
160 (FF01-536/40), a commercial SLR lens as the tube lens (Nikon Zoom-NIKKOR 80–200
161 mm f/4 AI-s), and a 5W LED (LZ1-00B200, 460 nm; LedEngin, San Jose CA). Data
162 acquisition was performed at 20 Hz, at a resolution of 1024 x 1024 pixels, with 16-bits
163 per pixel, for about 10-20 minutes. With 10x objective lens, the microscope provided a
164 field of view of 1.343 x 1.343 mm^2 (1.312 x 1.312 μm^2/pixel) of brain tissue. Imaging
165 data was streamed from the camera to RAM of a custom computer (dual Intel Xeon
166 processors, 128 GB RAM, and a GeForce GTX Titan video card) to ensure temporal
167 precision. After each imaging session, data was moved from RAM to hard drive and
168 saved in multi-page tagged image file format.

169 Two hippocampus datasets (A and B) were collected from two mice (dataset A was
170 previously reported by Mohammed et al. (Mohammed et al., 2016)). The mice were
171 trained to perform a trace conditioning task known to involve hippocampal neural activity
172 (Solomon et al., 1986, Moyer et al., 1990, Tseng et al., 2004, Sakamoto et al., 2005). In
173 this task, the animal was trained to associate a conditioned stimulus (a 350 ms long
174 tone) with an unconditioned stimulus (a gentle 100ms air puff to one eye). There was a

250 ms trace interval between two stimuli. During each recording session, the animal was head-fixed and performed 40 trials with a randomized 31-36 second inter-trial interval. The hippocampus datasets (1024 x 1024 pixels/frame, 2047 frames, ~100 seconds, ~4 GB size) analyzed in this study were part of larger recording sessions (~50GB size).

The striatum dataset was collected from a head-fixed animal running on a spherical treadmill system. The treadmill system consisted a styrofoam ball floated by air pressure in a 3D-printed bowl designed as described in (Dombeck et al., 2007) that allowed the animal to move its limbs freely while head-fixed. The mouse was first handled for several days before being head-fixed to the spherical treadmill. Habituation to running on the spherical treadmill while head-fixed occurred over 3-4 days/week at the same time of day as subsequent recording sessions (8-12 hours after lights-on), for several weeks. Single imaging sessions took approximately 25 minutes. Sampling occurred at approximately 20Hz and exposure time was fixed at 20ms. The striatum dataset (~100 seconds, ~4 GB size) contains 2047 frames with 1024 x 1024 pixels per frame and was also part of a larger dataset (~25GB size).

Two human referees manually identified ROIs in the hippocampus dataset A and in the striatum dataset to create a set of human-generated ROIs for comparison with ACSAT's segmentation results. This manual selection was done by viewing the image sequence and segmenting ROIs that had fluorescence traces compatible with neuronal dynamics and/or by selecting ROIs from a composite image created from the video sequence and confirming fluorescence traces were compatible with neuronal dynamics.

Wide-field cell culture dataset

The primary neuron cell dataset was collected from a 10-day-old neuron culture, infected with AAV9-Syn-GCaMP6.WPRE.SV40 virus. Seven days after infection, neurons were imaged at 20 Hz for 60 seconds. The primary neuron culture dataset contains 1201 frames, 1024 x 1024 pixels per frame, recorded with the same imaging setup as for the hippocampus and striatum datasets described above.

Two-photon dataset

The two-photon dataset was downloaded from the Neurofinder website (http://neurofinder.codeneuro.org/, 03.00). GCaMP6f was used as the indicator. The dataset contains 2250 frames with 498 x 490 pixels per frame with resolution 0.588 x 0.588 μm^2/pixel.

Signal-to-noise ratio (SNR) calculation

We calculated the signal-to-noise ratio (SNR) in decibels (dB) as

$$\text{SNR} = 20 \times \log \frac{\mu_{\text{ROI}}}{\sigma_{\text{background}}}$$

For the simulated datasets, μ_{ROI} is the mean intensity value of all pixels belonging to all ROIs in the time-collapsed image I_0, and similarly, $\sigma_{\text{background}}$ is the standard deviation of background pixel intensity values i.e. all pixels that do not belong to an ROI. For the hippocampus dataset A and the striatum dataset, μ_{ROI} is the maximum intensity value of an ROI trace, and $\sigma_{\text{background}}$ is the standard deviation of the background trace. The ROI trace value at each timepoint is the averaged intensity values of all pixels belonging to that ROI, and similarly for the background trace, which uses all pixels not belonging to

any ROI. Note that the SNR for the simulated datasets describes the whole time-collapsed image I_0, whereas the SNR for the hippocampus and striatum datasets describe an individual ROI.

Simulated datasets

We tested ACSAT's segmentation performance on 500 simulated datasets with varying signal-to-noise ratios (between ~19 dB and ~29 dB) and numbers of ROIs (between 300 and 700). Figure 2B shows some examples of the simulated time-collapsed image, i.e., the input image I_0 to ACSAT in Figure 1A. The simulation gives us the true locations of all ROIs so that we can accurately assess ACSAT's segmentation performance.

Our simulated datasets were obtained by a procedure adapted from (Zhou et al., 2018). We used the model $I_0 = E + AC$ to generate the simulated datasets where E represents noise, A represents the shapes of each ROI, and C adjusts each ROI's intensity to simulate uneven lighting.

The pixel noise values in E were randomly sampled from the background pixel values in the time-collapsed image for the hippocampus dataset. This noise is unlikely to be Gaussian because the time-collapsing procedure subtracts the mean value from the maximum value of each pixel such that the time-collapsed image is biased towards higher pixel values.

The centroid location of each ROI represented in A was randomly selected with weights C^2. The pixel values comprising the body of each ROI was modeled deterministically by

the bivariate Gaussian probability density function, with widths randomly selected according to Zhou et al. (Zhou et al., 2018).

The image C is also used to amplify each ROI's pixel values to reflect uneven lighting conditions across the imaging field. C was generated by applying heavy Gaussian filtering to the time-collapsed image of the hippocampus dataset until no individual ROIs are detectable.

Automated Cell Segmentation by Adaptive Thresholding (ACSAT) Overview

Fluorescence imaging data obtained in the form of image sequences is processed offline using a custom MATLAB algorithm. Image sequences were first motion-corrected as described in Mohammed et al. (2016) to remove micromotion of the imaged area caused by breathing and other movements of the animal. ACSAT (Figure 1A) is then applied to a time-collapsed image that represents the image sequences, to automatically identify individual neurons as regions of interest (ROIs).

The input image sequence is first loaded into MATLAB as a 3D matrix (height x width x time) and then time-collapsed to produce a representative two-dimensional image (height x width, I_0 in Figure 1A), where each pixel in I_0 is represented by the maximum intensity value of that pixel across the entire image sequence with its mean value removed. This time-collapsed image I_0 is then used for the rest of the algorithm. Pixels with low intensity values would correspond to static background, whereas pixels with high intensity values would correspond to neurons with GCaMP6 expression. In general, neurons with GCaMP6 expression appear in I_0 as a cluster of adjacent pixels with high intensity values and with size similar to that of a neuron. Meanwhile, it is improbable for

random background noise to generate clusters with similar properties. Thus, the time-collapsed image I_0 is expected to contain sufficient information to separate neurons from the background.

Next, ACSAT iteratively generates ROIs $\{ROIs\}_n$ from the time-collapsed image I_n for iterations $n = 1, 2, ...$, starting with $I_1 = I_0$. Prior to each subsequent iteration, I_n is generated by cumulatively clearing previously segmented ROIs, $\{ROIs\}_{n-1}$, from I_{n-1} by setting each ROI's pixels in I_{n-1} to blank values of 0 and dilating the cleared area. As described later, each iteration consists of both adaptive thresholding at the global level (Global FIBAT in Figure 1A), using the automatically selected threshold value τ_n^* (Figure 1B), and adaptive thresholding at the local level (Local FIBAT in Figure 1A). When the change in global threshold value $|\tau_{n+1}^* - \tau_n^*|$ is insignificant, further iterations are likely to contribute more false positives than true positives. Thus, the ACSAT algorithm terminates at iteration n if $\frac{|\tau_{n+1}^* - \tau_n^*|}{\tau_1^*} < \delta$ where τ_1^* acts as a normalizing factor. Accordingly, the final output of ACSAT is the union of the segmented ROIs from each iteration, $\{ROIs\}_1 \cup \cdots \cup \{ROIs\}_n$.

Global and Local Adaptive Thresholding in ACSAT

Each iteration n of ACSAT contributes a set of newly segmented ROIs $\{ROIs\}_n$ from I_n by applying our Fluorescence Intensity Based Adaptive Thresholding (FIBAT) algorithm, at the global and local levels (Global/Local FIBAT in Figure 1A). Briefly, FIBAT (Figure 1B) takes an inputted image I and outputs the optimal threshold value τ^* which results in optimally segmented ROIs $\{ROIs|\tau^*\}$.

Global adaptive thresholding is the first step in the nth iteration of ACSAT (Figure 1A). This step applies FIBAT directly to the whole image ($I_n \rightarrow I$) to identify potential ROIs ($\{ROIs|\tau^*\} \rightarrow \{ROIs\}'_n$).

These potential ROIs $\{ROIs\}'_n$ may include groups of adjacent neurons or overlapping neurons because neurons located above and below the focal plane could be captured in the same frame during wide-field imaging. Such overlap however is unlikely to occur in two-photon datasets or in cell culture datasets. The local adaptive thresholding step (Figure 1C) recursively separates any potentially overlapping ROIs within $\{ROIs\}'_n$ in order to output $\{ROIs\}_n$. Specifically, each ROI in $\{ROIs\}'_n$ is individually dilated and then inputted to the local FIBAT (ROI $\rightarrow I$) in (Figure 1B) to obtain a set of separated ROIs $\{ROIs|\tau^*\}$. If any outputted set $\{ROIs|\tau^*\}$ contains more than one separated ROIs, then each ROI in the set $\{ROIs|\tau^*\}$ is further separated using the same procedure, thus forming a recursive loop. Otherwise, if any outputted set $\{ROIs|\tau^*\}$ contains only one ROI, then the recursion terminates. The final output of the local thresholding step $\{ROIs\}_n$ is the union of all such sets containing one ROI that cannot be further separated.

Fluorescence Intensity Based Adaptive Thresholding (FIBAT)

As described, FIBAT (Figure 1B) is used in both the global and local adaptive thresholding steps of each iteration of ACSAT to identify potential ROIs in the time-collapsed image $I = I_n$ or to separate potentially overlapping neurons within I which is an element of $\{ROIs\}'_n$, respectively. In either case, an optimal pixel intensity threshold value τ^* separates ROIs from the background. FIBAT selects τ^* by searching for the

threshold value that maximizes the number of resulting ROIs that are larger in area than A_{min} and smaller in area than A_{max}.

The search is performed recursively over a pixel intensity range (τ_1, τ_T), where initially τ_1 is the minimum pixel intensity value in I and τ_T is the maximum pixel intensity value in I. From this search range, T test threshold values $\tau_1, \tau_2, \dots, \tau_{T-1}, \tau_T$ are uniformly selected. A larger T will decrease the probability of skipping the optimal threshold value, but it will result in more computation time that may not be necessary. Because the threshold value is refined by a recursive process until it reaches the optimal value that produces the maximum number of ROIs, the value of T should have little to no effect on ACSAT's segmentation results. We chose $T = 12$. Each of these test threshold values $\tau_1, \dots, \tau_T$ is applied to the image I by assigning each pixel a 1 (a true calcium event) if its value is greater than the threshold or a 0 (a false calcium event) otherwise. Morphological operations are then performed to refine the thresholded images. Specifically, these operations fill in holes (0s surrounded by 1s) and remove spur pixels which may be due to noise. The operations also break H-connected ROIs prior to splitting overlapping cells. ROIs are finally collected with 8-connectivity (MATLAB function bwlabel or bwconncomp) to generate a set of segmented ROIs for each test threshold value: $\{ROIs|\tau_1\}, \dots, \{ROIs|\tau_T\}$.

Since ROIs represent real neurons that are roughly spherical in shape and are about 5 μm - 20μm in diameter, some realistic criteria can be used to eliminate false ROIs that are not possibly actual neurons. Accordingly, FIBAT removes ROIs from $\{ROIs|\tau_1\}, \dots, \{ROIs|\tau_T\}$ if their centroid is outside the ROI, or if their area is less than

A_{min} or greater than A_{max}, or if their solidity (i.e. the area ratio between the convex hull of a ROI and the ROI itself) is greater than approximately the golden ratio.

The next search range is selected based on the results of the test thresholds. A relationship of the test threshold values $\tau_1, \dots, \tau_T$ vs the numbers of resulting ROIs $|\{ROIs|\tau_1\}|, \dots, |\{ROIs|\tau_T\}|$ can be generated (Figure 1B). If the test threshold value τ_c resulted in the most ROIs i.e. $c = \arg\max_c |\{ROIs|\tau_c\}|$, then the next search range is set to $(\tau_{\max\{1,c-1\}}, \tau_{\min\{T,c+1\}})$ in order to include τ_c inside the search range. If more than one test threshold value $\tau_{c_1}, \tau_{c_2}, \dots$ resulted in the same maximum number of ROIs, then the next search range is similarly set to $(\tau_{\max\{1,\min\{c_1,c_2,\dots\}-1\}}, \tau_{\min\{T,\max\{c_1,c_2,\dots\}+1\}})$ in order to contain all $\tau_{c_1}, \tau_{c_2}, \dots$. This search is terminated when further refinement of the search range produces little improvement in the number of detected ROIs: either the new search range $|\tau_{c+1} - \tau_{c-1}|$ is less than ϵ or the new range overlaps the previous range by at least α. We chose $\alpha = 90\%$ and $\epsilon =$ the smallest non-zero intensity difference between every pair of adjacent pixels in whole image I. As such, ϵ is determined automatically and does not require user input. Upon termination, the optimal threshold value is set to $\tau^* = \frac{1}{2}\left(\min\{\tau_{c_1}, \tau_{c_2}, \dots\} + \max\{\tau_{c_1}, \tau_{c_2}, \dots\}\right)$, and the segmented ROIs $\{ROIs|\tau^*\}$ includes ROIs whose area exceeds A_{max}.

Code accessibility

The code/software described in the paper is freely available online at [redacted for double-blind review]. The code is available in Extended Data 1.

Results

We tested ACSAT on 500 simulated datasets, two wide-field hippocampus datasets, a wide-field striatum dataset, a wide-field cell culture dataset, and a two-photon hippocampus dataset. The simulated datasets with known ground truth allowed us to accurately assess the segmentation performance of ACSAT in different conditions of SNR and number of ROIs. For the hippocampus dataset A and the striatum dataset, in which the ground truth is unknown, we used human-generated ROIs as a reference. For the cell culture dataset, hippocampus dataset B, and two-photon dataset, we provide the ACSAT segmented ROIs that can be inspected and interpreted by users.

ACSAT performance on simulated datasets with various SNRs and numbers of ROIs

To evaluate the performance of ACSAT, we simulated 500 time-collapsed images I_0 with various numbers of ROIs (between 300 and 700) at random locations and different SNRs (between ~19 dB and ~29 dB). The exact locations of ROIs are known and served as the ground truth to provide an accurate evaluation of the performance of ACSAT. For all 500 datasets, we used the parameters $\delta = 10\%$, $A_{min} = 50px \approx 86\mu m^2$ and $A_{max} = 300px \approx 516\mu m^2$ for the global adaptive thresholding step, and $A_{min} = 20px \approx 34\mu m^2$ and $A_{max} = \infty$ for the local adaptive thresholding step because ROIs tend to shrink in size after repeatedly applying FIBAT.

The recall and precision results for each of these simulated datasets are shown as dots in Figure 2A1 and Figure 2A2, respectively. Figure 2B shows examples of the simulated time-collapsed images, and each example corresponds to a dot in Figure 2A1 and Figure 2A2. At SNR greater than approximately 24 dB, ACSAT shows a stable performance with generally higher than 80% recall. The precision rate remains stable at

368 generally higher than 80% when SNR is greater than approximately 21 dB. However,
369 the performance of ACSAT falls when SNR is below approximately 20 dB.

370 ACSAT performance on hippocampus dataset A and striatum dataset

371 We used ACSAT (Figure 1A) to automatically segment ROIs from a hippocampus wide-
372 field imaging dataset and a striatum wide-field imaging dataset. Prior to the application
373 of the ACSAT, the image sequences were time-collapsed as shown in Figure 3 and
374 Figure 4 (top rows) for the hippocampus A and the striatum datasets, respectively.
375 These time-collapsed images show high intensity areas resembling neural morphology.
376 The final segmented ROIs outputted by ACSAT are illustrated in Figure 3 and Figure 4
377 (bottom row), respectively.

378 For both datasets, we initiated ACSAT using the same parameters as for the simulated
379 datasets ($\delta = 10\%$, $A_{min} = 50px \approx 86\mu m^2$ and $A_{max} = 300px \approx 516\mu m^2$ for the global
380 adaptive thresholding step, and $A_{min} = 20px \approx 34\mu m^2$ and $A_{max} = \infty$ for the local
381 adaptive thresholding step). To obtain the results as shown in Figure 3 and Figure 4, it
382 took approximately one minute per iteration on a Xeon E5-1650 v4 at 3.6GHz with 128
383 GB DDR4 RAM, but it used less than 30 MB RAM. As such, the RAM size had little
384 effect on the speed.

385 *ACSAT Performance Compared to Human-Generated ROIs*

386 To assess the performance of the ACSAT algorithm, we compared the ACSAT
387 segmentation results with ROIs generated by human inspection (human-generated
388 ROIs). This set of human-generated ROIs contained 423 ROIs for the hippocampus
389 dataset A and 91 ROIs for the striatum dataset. We first compared the ACSAT-

generated ROIs for the hippocampus A and striatum datasets with the ROIs in the human-generated ROIs. We consider a pair of ROIs to correspond to the same neuron if they had centroids that were less than $50px \approx 65.6\mu m$ apart and had a mutual overlap greater than 60%. We calculated the mutual overlap as the average of the percentages of the overlapping area against the areas of both ROIs. When there were multiple ROIs sharing overlapping areas, we selected the pair with highest mutual overlap as the matched ROIs.

For the hippocampus dataset A, ACSAT identified 445 ROIs after three iterations. Among these 445 ROIs, 317 ROIs were matched in the human-generated ROIs (Match), and 128 ROIs were not in the human-generated ROIs (ACSAT-only). Additionally, 106 ROIs in human-generated ROIs were not identified by ACSAT (Human-only). This result gave us a precision rate of 71.2% (317 out of 445) and a recall rate of 74.9% (317 out of 423). For the striatum dataset, ACSAT was terminated after one iteration and identified a total of 135 ROIs: 69 Match ROIs, 66 ACSAT-only ROIs, and 22 Human-only ROIs (precision rate: 51.1%, recall rate: 75.8%).

We further examined the fluorescence traces of ROIs from the ACSAT-only, Human-only, and Match groups. Representative traces are shown in Figure 5A1 and B1, respectively for the hippocampus A and striatum datasets, and all traces are available in Extended Data 2. The value of each ROI fluorescence trace at each timepoint is the average intensity value of all pixels belonging to that ROI. In Figure 5A1 and B1, each trace is normalized by subtracting the mean value of that trace over time and then dividing this difference by that mean value. We calculated the signal-to-noise ratio

412 (SNR) for every ROI in each group. In both the hippocampus A and striatum datasets,
413 the Match ROIs exhibit a broad range of SNR, indicating that both ACSAT and humans
414 are capable of identifying ROIs with various intensities in the time-collapsed image
415 (Figure 5A2 and B2).

416 We further examined the individual ROIs identified by ACSAT that were not identified by
417 humans. This secondary manual inspection found that some of the ACSAT-only ROIs
418 were actually true neurons (i.e. with fluorescence traces compatible with neuronal
419 dynamics) that were missed in the initial human-generated ROIs due to human error.
420 This means that ACSAT was able to segment ROIs that were difficult to identify by
421 human experts. Specifically, for the hippocampus A dataset, 70 (54.7%) out of 128
422 ROIs initially labeled as ACSAT-only were later determined to be actual neurons, and
423 for the striatum dataset, 31 (47%) ROIs were true neurons. After correction, of the total
424 445 ACSAT ROIs from the hippocampus dataset A, 387 segmented ROIs corresponded
425 to true neurons (Match), and 58 segmented ROIs did not correspond to true neurons
426 determined by human inspection (ACSAT-only). Additionally, 106 true ROIs were not
427 segmented (Human-only). This corresponds to a precision rate of 87% and a recall rate
428 of 78.5%. Similarly, for the striatum dataset, which resulted in 135 ACSAT ROIs, there
429 were 100 Match ROIs, 35 ACSAT-only ROIs, and 22 Human-only ROIs after correction.
430 This corresponds to a precision rate of 74.1% and a recall rate of 82%. Even though
431 neurons in the hippocampus and striatum have different morphology and fluorescence
432 intensity, ACSAT was consistently effective for both datasets, and it was able to detect
433 low-intensity neurons that were initially undetected by human referees. As such, our
434 results demonstrate the robustness and effectiveness of the algorithm.

435 The result from the hippocampus dataset A shows that ACSAT successfully identified
436 true ROIs of diverse sizes (Figure 6, red). In general, the false positive ROIs had
437 relatively smaller areas (Figure 6, yellow), similar to the ROIs missed by human
438 referees (Figure 6, green). This indicates that ACSAT is more likely to recognize
439 intensity changes in small areas thereby outperforming human referees under such
440 challenging detection conditions. Additionally, ACSAT missed a small portion of true
441 ROIs, which shares similar sizes with those identified (Figure 6, blue).

442 *Number of Iterations in Using ACSAT*

443 For the hippocampus dataset A, ACSAT was terminated at iteration $n = 3$ when the
444 change in global threshold value $\frac{|\tau_4^* - \tau_3^*|}{\tau_1^*} = 5.12\% < \delta = 10\%$. For the striatum dataset,
445 ACSAT was terminated at iteration $n = 1$ when the change in global threshold value
446 $\frac{|\tau_2^* - \tau_1^*|}{\tau_1^*} = 4.53\% < \delta = 10\%$.

447 To evaluate how ACSAT performs when terminated at different iteration numbers, we
448 ran ACSAT up to 9 iterations on both datasets, and calculated several major
449 performance indicators after each iteration (Figure 7): cumulative number of ROIs,
450 global threshold value, recall, false negative rate, and false discovery rate (which is
451 equal to 1 – precision) compared to the human-generated ROIs prior to secondary
452 manual inspection of false positives. The cumulative number of ROIs, recall, and false
453 discovery rate increased with the iteration number, but at different speed. While the
454 cumulative number of ROIs and the false discovery rate increased steadily, recall rose
455 steeply and reached its plateau within approximately three iterations for the
456 hippocampus dataset and after the first iteration for the striatum dataset. Both the global

threshold value and the false negative rate dropped as iterations progressed, indicating ACSAT dynamically adjusted the threshold to capture potential ROIs with lower intensity in later iterations. This dynamic adjustment of the threshold value at each iteration was only possible because of the removal of segmented ROIs prior to each iteration. Overall, the changes in these performance indicators over iterations suggested that most true ROIs were identified during the early iterations: $n \leq 3$ for the hippocampus dataset and $n = 1$ for the striatum dataset, which are consistent with when the ACSAT termination criterion described by δ was met. ROIs segmented during later iterations were mostly false positive.

FIBAT Global and Local Thresholding

In Figure 8, we demonstrate how FIBAT (Figure 1B) determines the threshold value that achieves optimal segmentation results by sampling the distribution of threshold values vs the number of ROIs. Each trace of Figure 8 plots the number of ROIs that results from each sampled threshold value in the global thresholding step during the first four iterations of ACSAT (Figure 1A) on the hippocampus dataset. In each iteration, FIBAT (Figure 1B) first samples the threshold values across the entire intensity range at coarse resolution to identify the potential search range that may result in the maximum number of ROIs. FIBAT further re-samples threshold values within the new search range with a finer resolution, until it reaches a threshold value that gives the maximum number of ROIs. This design allows FIBAT to determine the optimal threshold value with a fine resolution without actually sampling the whole intensity range at the fine scale, and, as a result, reduces the processing time.

After performing global thresholding to identify potential ROIs $\{ROIs\}'_n$ (Figure 1A), ACSAT further applies FIBAT locally to each identified ROI in $\{ROIs\}'_n$ to refine the segmentation results (Figure 9). When neurons are densely labeled with GCaMP6, using the global thresholding step alone may lead to one or more large clusters of adjacent neurons being segmented as a single ROI (Figure 9A). For each such cluster, FIBAT (Figure 1B) determines and applies a new threshold value to the local ROI area. With local thresholding, the example cluster is further segmented into 5 new ROIs (Figure 9B), which would not otherwise be separated by applying the global threshold. Because further local thresholding produces the same result (Figure 9C), the local thresholding step of ACSAT concludes that these 5 ROIs cannot be further separated, exits the recursive loop, and outputs these ROIs.

To evaluate the efficacy of local thresholding, we examined the hippocampus dataset A at each iteration before and after the local thresholding step (Figure 10, left and right bars, respectively). Local thresholding refined the ROIs detected by global thresholding and captured more true ROIs at every iteration. It is also worth noting that, at later iterations, local thresholding was still able to identify true ROIs that were missed by global thresholding alone (Figure 10, iteration 4).

ACSAT performance on two-photon dataset

We applied ACSAT to the two-photon dataset Neurofinder 03.00 (Figure 11C). Genetically Encoded Calcium Indicators are generally not expressed in the nuclei (Tian et al., 2009), and because of the optical sectioning technique that two-photon imaging provides, in this dataset the nuclei appear dark. Additionally, this dataset had high speckle noise. Thus, the time-collapsed image generated by ACSAT using max minus

mean pixel values shows bright nuclei. The truth file provided by Neurofinder contains 621 ROIs, most of which are nuclei. Since the features of this dataset are the nuclei, which are smaller, we used the parameters $\delta = 5\%$, $A_{min} = 20px \approx 6.9\mu m^2$ and $A_{max} = 1000px \approx 34.6\mu m^2$ for the global adaptive thresholding step, and $A_{min} = 20px \approx 6.9\mu m^2$ and $A_{max} = \infty$ for the local adaptive thresholding step.

ACSAT identified 571 ROIs. Among these, 442 ROIs were matched with the truth (true positive), and 179 ROIs were not in the truth (false positive). Additionally, 129 ROIs in truth were not identified by ACSAT (false negative). This result gave us a recall rate of 71.2% (442 out of 621) and a precision rate of 77.4% (442 out of 571).

We further inspected the time-collapsed image I_0 and observed that the right side of the time-collapsed image I_0 had different patterns of texture than the left side. To utilize the new texture information for ROI detection by ACSAT, we extracted the right side of I_0 that is rich in texture information to generate I_0' as input to ACSAT. The I_0' was generated by change detection between the original image and its Gaussian-filtered counterpart. Thus, ACSAT identified an additional 157 ROIs, of which 95 were true positives, and 62 were false positives. Combining these additional ROIs with the ROIs identified by direct application of ACSAT results in a recall rate of 82.8% (514 out of 621) and a precision rate of 70.6% (514 out of 728).

ACSAT performance on cell culture and hippocampus B dataset

Finally, we used ACSAT to detect ROIs in the dataset of the primary neuron culture expressing GCaMP6f (Figure 11B). Qualitatively, it appears ACSAT successfully identified the cell bodies of the majority of neurons in early iterations, and neurites in later iterations. We also used ACSAT to detect ROIs in the hippocampus dataset B

(Figure 11A). For both datasets, we used the parameters $\delta = 10\%$, $A_{min} = 50px \approx 86\mu m^2$ and $A_{max} = 300px \approx 516\mu m^2$ for the global adaptive thresholding step, and $A_{min} = 20px \approx 34\mu m^2$ and $A_{max} = \infty$ for the local adaptive thresholding step because ROIs tend to shrink in size after repeatedly applying FIBAT.

Discussion

In this study, we presented our Automated Cell Segmentation by Adaptive Thresholding (ACSAT) method that adaptively selects threshold values based on image pixel intensity with two iterative steps at the global and local levels using a time-collapsed image. As such, the algorithm is capable of handling morphological variations in fluorescence intensity in neurons and is robust against luminance condition changes across datasets. When applied to two datasets collected from the hippocampus and the striatum in mice, ACSAT resulted in approximately 80% recall rate of regions of interest (ROIs) containing individual neurons (78.5% for the hippocampus A dataset and 82% for the striatum dataset), and approximately 80% precision rate (87% for the hippocampus dataset and 74.1% for the striatum dataset). ACSAT was also able to detect low-intensity ROIs that were initially undetected by human referees. When applied to 500 simulated datasets, ACSAT achieved recall and precision rates higher than 80% when SNR was no less than approximately 24 dB. However, the performance of ACSAT falls when SNR reaches below approximately 20 dB.

The ACSAT algorithm is an intuitive thresholding method that uses global and local schemes to address variations in fluorescence intensity levels of GCaMP6 fluorescence even within the same image field. Simply applying a lower global threshold value would result in few large ROIs containing multiple neurons within one ROI. On the other hand,

548 with a high global threshold value, only a small number of neurons with high intensity
549 would be found. As such, applying a single high or low threshold value would generate
550 inadequate results of either few or excessive ROIs, which is a universal limitation of
551 thresholding methods. Our algorithm efficiently addresses this challenge in two ways.

552 First, it cumulatively excludes previously segmented ROIs from the time-collapsed
553 image I_n after each iteration so that in the following iteration, ACSAT could detect new
554 ROIs that require distinct thresholds to separate but were missed with previous
555 thresholds. Therefore, the global threshold value τ_n^* (Figure 1) used by ACSAT usually
556 decreases after each iteration, and ROIs with high intensity were segmented before
557 those with low intensity, as shown in Figure 3 and Figure 4. Because ACSAT is based
558 on adaptive thresholding, it allows us to objectively and robustly segment ROIs with low
559 intensity relative to the background. These low-intensity areas often pose challenges to
560 human experts when manually detecting ROIs, as our results showed that about half of
561 the ROIs initially labeled as false positive were actually true neurons (Figure 10).

562 Second, ACSAT uses Fluorescence Intensity Based Adaptive Thresholding (FIBAT)
563 locally to separate overlapping ROIs. This approach directly addresses the issue of
564 heterogeneity in recorded neural signals when the intensities of pixels surrounding an
565 ROI can vary. However, because a higher thresholding value is usually required to
566 separate adjoining neurons, the output sub-ROIs after local FIBAT are often smaller
567 than the corresponding true neurons. Thus, a simple dilation step was applied during
568 the local FIBAT step. This correction is useful to prevent real ROIs from falling below
569 the minimum area criteria A_{min} and thus being removed. Although the interleaving
570 process of global FIBAT and local FIBAT has been effective in addressing overlapping

neurons, a potential problem still arises if two neurons with similar intensities have significant overlap with each other in the time-collapsed image such that there is no trough between them. Then ACSAT may identify them as a single ROI. Conversely, if there is a neuron with multiple hot-spots (Pnevmatikakis et al., 2016), then this may be identified as multiple neurons by ACSAT. Such a scenario, however, can be minimized by the minimum area criteria A_{min} and the maximum area criteria A_{max}. Spatial overlap is profound for wide-field imaging, but not for two-photon imaging or in vitro cell culture imaging with single cell layer. With increasing improvement wide-field imaging, such as volumetric imaging (Shain et al., 2018, Xiao et al., 2018), such significant overlap may be better eliminated during data acquisition step.

ACSAT has three sets of free parameters that can be rationally chosen or otherwise are not sensitive: δ which describes the termination condition for ACSAT, α which describes a termination condition for FIBAT, and A_{min} and A_{max} which describe the allowed sizes of ROIs.

The termination condition for ACSAT, described by δ, can be explained by the tendencies of ACSAT. Specifically, running ACSAT for more iterations increases the number of ROIs segmented, especially the number of low-intensity ROIs, as the global threshold value τ_n^* gradually decreases (Figure 7). While many of the added ROIs are true ROIs, the proportion of false positive ROIs added increases as iteration number increases (Figure 7). This increasing proportion of outputted false positives in later iterations can be attributed to the higher probability of a spurious collection of adjacent background pixels meeting the criteria to be an ROI. Also, the added false positives can be related to the step which clears previously segmented ROIs from the time-collapsed

image at the start of each iteration of ACSAT. Due to the scattering of light in brain tissue, ROI removal may leave a few small fragments of bright pixels around removed areas, which could be identified as ROIs during the next iteration. ACSAT tries to avoid this problem by dilating the cleared area which makes sure the whole ROI is cleared rather than only the brighter center. Besides dilation, these misidentified ROIs were also discarded either because of their small size or because they do not meet the solidity criteria; however, occasionally they may pass the size criteria and become the false positive ROIs. As a result, the majority of false positives tend to have small size (Figure 6, yellow).

In order to balance the effects of simultaneous increase in true ROIs and false positive ROIs, ACSAT stops when a decrease of global threshold value becomes relatively small between iterations i.e. $\frac{|\tau_{n+1}^* - \tau_n^*|}{\tau_1^*} < \delta$. At that stage, most true ROIs have been detected and removed from the time-collapsed image. Thus, the global threshold values τ_n^* of any further iterations are similar, so most ROIs detected at this stage are false positives. For the hippocampus dataset A, iteration $n = 3$ is when the increase in false positives begins to outweigh the increase in true positives, and for the striatum dataset, nearly all true ROIs segmented by ACSAT were outputted at iteration $n = 1$ (Figure 7). Qualitatively, the time-collapsed image I_0 for hippocampus has a higher density of neurons with a greater variety of pixel intensities than the I_0 for striatum, so it may take more iterations for ACSAT to perform at the same rate on the hippocampus dataset than on the striatum dataset. ACSAT's performance under the diverse conditions of these two datasets suggests that our choice of $\delta = 10\%$ provides a robust and rational termination condition for ACSAT that can be generalized to other datasets, namely the

500 simulated datasets and the cell culture dataset, as well. In fact, changing the termination condition from $\delta = 10\%$ to $\delta = 5\%$ only affected the segmentation results in less than 17% of the 500 simulated datasets. For the two-photon dataset, our reported results are using the termination condition $\delta = 5\%$. In general, users can choose δ to be between 5% and 10% based on the needs of their application: if recall is more important, then users should choose a smaller δ, and if precision is more important, then users should choose a larger δ.

Additionally, the final segmentation results generated by ACSAT are not sensitive to the termination conditions for FIBAT described by α and ϵ. FIBAT is terminated if the threshold search range has minimal change over an iteration, which we determine in two ways. One way this condition would be satisfied is when all threshold values within the search range result in the same, optimal number of ROIs. This is equivalent to setting the criteria $\alpha = 100\%$. For the practical purpose of reducing FIBAT run time, we allow termination if the change in the search range is less than $1 - \alpha = 10\%$. This condition is also easily met when FIBAT is used in the local thresholding step because, by definition, ROIs that cannot be separated by FIBAT will return exactly one ROI no matter what threshold value is used. Additionally, we terminate FIBAT if the search range is smaller than ϵ, the smallest difference between any pair of adjacent pixels in I, which can be objectively and automatically determined from I. If FIBAT were to continue refining the threshold value, then the gained precision beyond that defined by ϵ would be useless due to the discrete step in pixel intensity values in I.

The last set of parameters A_{min} and A_{max} should be chosen based on how large neurons are expected to be using information including neuron size, image resolution,

magnification, imaging method, etc. In our wide-field datasets, the boundaries of neurons may not be as well-defined as those collected with two-photon microscope, and the size will appear larger than the size of a neuron due to light scattering in wide-field conditions. This effect is consistent with our observation that the minimum size of the human-generated ROIs was $38px \approx 64.6\mu m^2$ for the hippocampus A dataset and $66px \approx 112.2\mu m^2$ for the striatum dataset. Thus, our minimum ROI criteria for the wide-field datasets may be larger than a typical neuron size.

The images I_0 used by ACSAT are time-collapsed, and therefore do not contain any temporal information. With the flexibility of ACSAT, the framework of ACSAT can be used as long as a single image can be generated to represent the ROIs within the image sequence. For example, an input image I_0' can be generated where the value of each pixel represents the time of its maximum intensity. This image I_0' would allow ACSAT to separate adjoined ROIs that have similar intensity values in I_0 but reach their maximum intensity at different time points, which is described by I_0'. Other ways to generate the single representative image include correlations with nearby pixels, intensity dynamics such as standard deviation or variance over time, texture of the time-collapsed image (for example, as used for the two-photon dataset), and a combination of various parameters. Overall, by taking advantage of adaptively determining the threshold value at both the global level and the local level, ACSAT can theoretically perform segmentation on any image containing ROIs with non-homogenous intensity as long as it has sufficient contrast between ROIs and the background.

References

Allen WE, Kauvar IV, Chen MZ, Richman EB, Yang SJ, Chan K, Gradinaru V, Deverman BE, Luo L, Deisseroth K (2017) Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. Neuron 94:891-907 e896.

Andermann ML, Kerlin AM, Reid RC (2010) Chronic cellular imaging of mouse visual cortex during operant behavior and passive viewing. Front Cell Neurosci 4:3.

Chen TW, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, Looger LL, Svoboda K, Kim DS (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499:295-300.

Dombeck DA, Khabbaz AN, Collman F, Adelman TL, Tank DW (2007) Imaging large-scale neural activity with cellular resolution in awake, mobile mice. Neuron 56:43-57.

Fantuzzo JA, Mirabella VR, Hamod AH, Hart RP, Zahn JD, Pang ZP (2017) Intellicount: High-Throughput Quantification of Fluorescent Synaptic Protein Puncta by Machine Learning. eNeuro 4.

Ghosh KK, Burns LD, Cocker ED, Nimmerjahn A, Ziv Y, Gamal AE, Schnitzer MJ (2011) Miniaturized integration of a fluorescence microscope. Nat Methods 8:871-878.

Hamel EJ, Grewe BF, Parker JG, Schnitzer MJ (2015) Cellular level brain imaging in behaving mammals: an engineering approach. Neuron 86:140-159.

Hofer SB, Ko H, Pichler B, Vogelstein J, Ros H, Zeng H, Lein E, Lesica NA, Mrsic-Flogel TD (2011) Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. Nat Neurosci 14:1045-1052.

Huber D, Gutnisky DA, Peron S, O'Connor DH, Wiegert JS, Tian L, Oertner TG, Looger LL, Svoboda K (2012) Multiple dynamic representations in the motor cortex during sensorimotor learning. Nature 484:473-478.

Issa JB, Haeffele BD, Agarwal A, Bergles DE, Young ED, Yue DT (2014) Multiscale optical Ca2+ imaging of tonal organization in mouse auditory cortex. Neuron 83:944-959.

Kim TH, Zhang Y, Lecoq J, Jung JC, Li J, Zeng H, Niell CM, Schnitzer MJ (2016) Long-Term Optical Access to an Estimated One Million Neurons in the Live Mouse Cortex. Cell reports 17:3385-3394.

Lutcke H, Gerhard F, Zenke F, Gerstner W, Helmchen F (2013) Inference of neuronal network spike dynamics and topology from calcium imaging data. Front Neural Circuits 7:201.

Mellen NM, Tuong CM (2009) Semi-automated region of interest generation for the analysis of optically recorded neuronal activity. Neuroimage 47:1331-1340.

Mohammed AI, Gritton HJ, Tseng HA, Bucklin ME, Yao Z, Han X (2016) An integrative approach for analyzing hundreds of neurons in task performing mice using wide-field calcium imaging. Sci Rep 6:20986.

Moyer JR, Jr., Deyo RA, Disterhoft JF (1990) Hippocampectomy disrupts trace eye-blink conditioning in rabbits. Behav Neurosci 104:243-252.

Mukamel EA, Nimmerjahn A, Schnitzer MJ (2009) Automated analysis of cellular signals from large-scale calcium imaging data. Neuron 63:747-760.

Ohki K, Chung S, Ch'ng YH, Kara P, Reid RC (2005) Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. Nature 433:597-603.

Otsu N (1979) Threshold Selection Method from Gray-Level Histograms. Ieee T Syst Man Cyb 9:62-66.

Pinto L, Dan Y (2015) Cell-Type-Specific Activity in Prefrontal Cortex during Goal-Directed Behavior. Neuron 87:437-450.

Pnevmatikakis EA, Soudry D, Gao Y, Machado TA, Merel J, Pfau D, Reardon T, Mu Y, Lacefield C, Yang W, Ahrens M, Bruno R, Jessell TM, Peterka DS, Yuste R, Paninski L (2016) Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. Neuron 89:285-299.

712 Poort J, Khan AG, Pachitariu M, Nemri A, Orsolic I, Krupic J, Bauza M, Sahani M, Keller GB,
713 Mrsic-Flogel TD, Hofer SB (2015) Learning Enhances Sensory and Multiple Non-
714 sensory Representations in Primary Visual Cortex. Neuron 86:1478-1490.
715 Sadeghian F, Seman Z, Ramli AR, Abdul Kahar BH, Saripan MI (2009) A framework for white
716 blood cell segmentation in microscopic blood images using digital image processing. Biol
717 Proced Online 11:196-206.
718 Sakamoto T, Takatsuki K, Kawahara S, Kirino Y, Niki H, Mishina M (2005) Role of hippocampal
719 NMDA receptors in trace eyeblink conditioning. Brain Res 1039:130-136.
720 Sezgin M, Sankur B (2004) Survey over image thresholding techniques and quantitative
721 performance evaluation. J Electron Imaging 13:146-168.
722 Shain WJ, Vickers NA, Li J, Han X, Bifano T, Mertz J (2018) Axial localization with modulated-
723 illumination extended-depth-of-field microscopy. Biomed Opt Express 9:1771-1782.
724 Shen S, Syal K, Tao N, Wang S (2015) An automated image analysis method for high-
725 throughput classification of surface-bound bacterial cell motions. Rev Sci Instrum
726 86:126104.
727 Solomon PR, Vander Schaaf ER, Thompson RF, Weisz DJ (1986) Hippocampus and trace
728 conditioning of the rabbit's classically conditioned nictitating membrane response. Behav
729 Neurosci 100:729-744.
730 Sun XR, Badura A, Pacheco DA, Lynch LA, Schneider ER, Taylor MP, Hogue IB, Enquist LW,
731 Murthy M, Wang SS (2013) Fast GCaMPs for improved tracking of neuronal activity. Nat
732 Commun 4:2170.
733 Tian L, Hires SA, Mao T, Huber D, Chiappe ME, Chalasani SH, Petreanu L, Akerboom J,
734 McKinney SA, Schreiter ER, Bargmann CI, Jayaraman V, Svoboda K, Looger LL (2009)
735 Imaging neural activity in worms, flies and mice with improved GCaMP calcium
736 indicators. Nat Methods 6:875-881.
737 Tseng W, Guan R, Disterhoft JF, Weiss C (2004) Trace eyeblink conditioning is hippocampally
738 dependent in mice. Hippocampus 14:58-65.
739 Wachowiak M, Economo MN, Diaz-Quesada M, Brunert D, Wesson DW, White JA, Rothermel
740 M (2013) Optical dissection of odor information processing in vivo using GCaMPs
741 expressed in specified cell types of the olfactory bulb. J Neurosci 33:5285-5300.
742 Wilt BA, Fitzgerald JE, Schnitzer MJ (2013) Photon shot noise limits on optical detection of
743 neuronal spikes and estimation of spike timing. Biophys J 104:51-62.
744 Xiao S, Tseng HA, Gritton H, Han X, Mertz J (2018) Video-rate volumetric neuronal imaging
745 using 3D targeted illumination. Sci Rep 8:7921.
746 Zhou P, Resendez SL, Rodriguez-Romaguera J, Jimenez JC, Neufeld SQ, Giovannucci A,
747 Friedrich J, Pnevmatikakis EA, Stuber GD, Hen R, Kheirbek MA, Sabatini BL, Kass RE,
748 Paninski L (2018) Efficient and accurate extraction of in vivo calcium signals from
749 microendoscopic video data. Elife 7.
750 Ziv Y, Burns LD, Cocker ED, Hamel EO, Ghosh KK, Kitch LJ, El Gamal A, Schnitzer MJ (2013)
751 Long-term dynamics of CA1 hippocampal place codes. Nat Neurosci 16:264-266.

752

753

Figure Captions

Figure 1. **Flowchart of the Automated Cell Segmentation by Adaptive Thresholding (ACSAT) algorithm.** (A) The input is the time-collapsed image I_0, and the output is a collection of automatically segmented ROIs. In each iteration, the "Global FIBAT" step identifies potential ROIs $\{ROIs\}'_n$ by applying FIBAT, described in part B, to the entire image I_n; and the "Local FIBAT" step, described in part C, splits overlapping ROIs. (B) Flowchart of the Fluorescence Intensity Based Adaptive Thresholding (FIBAT) algorithm. The input image is segmented using each of the test threshold values $\tau_1, \tau_2, \ldots, \tau_T$. The search range for the optimal threshold value (τ_1, τ_T) is iteratively narrowed to contain the test threshold value which results in the maximum number of ROIs. (C) Local FIBAT procedure. FIBAT, described in part B, is recursively applied to each potential ROI until the resulting ROI(s) can no longer be separated by FIBAT.

Figure 2. **ACSAT performance on simulated datasets**. (A1) Recall and (A2) precision are plotted as a function of signal-to-noise ratio (SNR) and number of ROIs. Each dot corresponds to the ACSAT result for one of the 500 simulated datasets. A surface was fitted to these dots for visualization. The black vertical plane corresponds to the SNR of the hippocampus A dataset. (B) Six examples of simulated time-collapsed images, labeled a-f, correspond to the dots labeled a-f in parts A1 and A2.

Figure 3. **Hippocampus dataset A and ROIs identified by ACSAT.** (A) The aggregated image of hippocampus dataset A and zoom-in images (Ai, Aii, and Aiii, corresponding to the grey boxes). (B) ACSAT determined ROIs from multiple iterations overlying on the aggregated image (red, yellow, green, and blue outline corresponds to

776 iteration 1, 2, 3, and 4, respectively). The fourth iteration (blue) is shown for comparison
777 although ACSAT terminated at iteration 3 (red, yellow, and green).

778 Figure 4. **Striatum dataset and ROIs identified by ACSAT.** (A) The aggregated image
779 of striatum dataset and zoom-in images (Ai, Aii, and Aiii, corresponding to the grey
780 boxes). (B) ACSAT determined ROIs from multiple iterations overlying on the
781 aggregated image (red, yellow, green, and blue outline corresponds to iteration 1, 2, 3,
782 and 4, respectively). The second (yellow), third (green), and fourth (blue) iterations are
783 shown for comparison although ACSAT terminated at iteration 1 (red).

784 Figure 5. **Fluorescence traces and signal-to-noise ratios**. (A1) Representative
785 fluorescence traces from the hippocampus dataset A for ROIs identified by both ACSAT
786 and human referees (Match). ROIs identified only by ACSAT (ACSAT), and ROIs
787 identified only by human referees (Human). (A2) Histogram of signal-to-noise ratio
788 (SNR) for Match, ACSAT, and Human ROIs from the hippocampus dataset A. (B1)
789 Representative fluorescence traces from the striatum dataset. (B2) Histogram of SNR
790 for the striatum dataset.

791 Figure 6. **Distribution of ROI size.** ACSAT identified true ROIs (red) with various size.
792 The false positive ROIs (yellow) and those missed by human experts (green) tend to
793 have small areas, while the areas of false negative ROIs (blue) appear slightly larger.

794 Figure 7. **Performance of ACSAT over iterations.** For both (A) hippocampus dataset
795 A and (B) striatum dataset, the cumulative number of identified ROIs (blue line)
796 increased steadily over iterations. The global threshold (dash blue line) tended to
797 decrease with each iteration, allowing ACSAT to capture ROIs with lower intensity. Both

recall (solid red line) and false discovery rate = 1 – precision (dotted red line) increased with iterations, while the false negative rate (dashed red line) decreased. All results reported here are based on human-generated ROIs prior to secondary manual inspection of false positives.

Figure 8. **Convergence of the FIBAT optimal global threshold value for the hippocampus dataset A.** FIBAT first sampled at a coarse scale across a wide intensity range, and then focused on a small potential intensity range with a fine scale to identify the optimal global threshold value that generated most ROIs. The vertical lines indicate the final optimal global threshold value determined by FIBAT for each iteration.

Figure 9. **Improved ROI identification by local thresholding.** (A) With global thresholding alone, a cluster of hippocampal neurons was identified as a single ROI. (B) After application of local thresholding, ACSAT successfully separated five new ROIs from the single ROI. (C) Zoom-in of each ROI separated by local thresholding.

Figure 10. **Local thresholding improves ACSAT performance.** The ROIs identified by ACSAT at each iteration before local thresholding (left) and after (right). Local thresholding separated overlapping ROIs and thus helped identify more ROIs, including those identified by human (black) or missed by human (red).

Figure 11. **ACSAT results of various datasets**. (A) The time-collapsed image of hippocampus dataset B (top) with ACSAT ROIs overlaid (bottom). (B) The time-collapsed image for the primary neuron culture dataset (top) with ACSAT ROIs overlaid (bottom). (C) The time-collapsed image for the two-photon dataset (Neurofinder 03.00) (top) with ACSAT ROIs overlaid (bottom). For all three datasets, the majority of ROIs

were identified during the first two iterations. Red, yellow, green, and blue ROI outline corresponds to iteration 1, 2, 3, and 4, respectively.

Extended Data 1. ZIP file contains eleven MATLAB files which comprise the ACSAT algorithm.

Extended Data 2. Fluorescence traces in a MATLAB struct for all ROIs in the hippocampus dataset A and the striatum dataset grouped into “Match,” “ACSAT,” and “Human” as defined in Figure 5. Groupings here are based on human-generated ROIs prior to secondary manual inspection of ACSAT-only ROIs.