Multi-scale automated cell segmentation in two-photon calcium imaging

Bronwyn Woods[13], Alberto Vazquez[23], Seong-Gi Kim[23], and William Eddy[13]

[1] Department of Statistics, Carnegie Mellon University, Pittsburgh PA
[2] Department of Radiology, University of Pittsburgh, Pittsburgh PA
[3] Center for the Neural Basis of Cognition, Pittsburgh PA

Abstract. Two photon calcium imaging (TPCI) is a relatively new and very promising technique for *in vivo* imaging of the structure and function of neural populations. However, the data processing methodology for TPCI is underdeveloped. This presents an opportunity for statistics and machine learning to contribute substantively to basic neuroscience by providing a principled analysis pipeline that can be used by experimenters. We present here a procedure for automating the detection of cells in TPC images. Our procedure consists of an unsupervised multi-scale blob detector to generate candidate cell masks, and a minimally supervised Random Forest classifier to verify or discard the candidates. Our procedure improves on existing techniques, requiring minimal supervision, adapting to spatial inhomogeneities, and applying generally over cell types and animals.

1 Introduction

Two-photon calcium imaging (TPCI) is an increasingly popular *in vivo* fluorescence imaging technique [1–4]. Through the use of a two photon laser scanning microscope to image tissue containing a calcium responsive functional indicator as well as a static structural dye, TPCI can reveal the location, size, shape, and activity of neurons, astrocytes and blood vessels. From this, experimenters can deduce neural spike trains [5,6], calcium transients in astrocytes [7–9], properties of local blood flow [10], and connectivity of the local neural network [11]. TPCI gives more comprehensive measurements of cortical function than other recording techniques with similar spatial scale and is therefore ideal for studying integrative questions in basic neuroscience.

Despite the promise of TPCI for experimental neuroscience, appropriate processing methodology is still in its infancy. Only recently have problems started to be addressed such as motion correction [12–14], segmentation of regions of interest [15–18], integration of fluorescence over

regions of interest [13], and inference of spike trains [1, 2, 5, 6]. Here we address the second of these problems: the segmentation of TPCI images to identify cells.

Typically, TPCI can be viewed as videos of fluorescence over time in a particular plane of cortex. The current standard is to segment these images roughly by hand or with ad-hoc semi-manual tools. This is tedious for the experimenter, and not standardizable between labs. Automated segmentation will reduce the tedium of manual annotation, and may also increase precision by using a richer representation of the data than can be provided to a human.

Two groups have recently proposed automated techniques for this task. Mukamel et. al. [17] proposed a procedure based on spatio-temporal ICA. Their approach relies strongly on sparse fluorescence transients in order to detect a cell and can therefore only detect active neurons. Valmianski and colleagues [16] proposed a two-stage supervised classifier. They first use data manually annotated on the pixel level to classify pixels as part of a cell or not, and then classify groups of pixels as either true cells or false positives. This approach has the advantage of being able to segment any cell, but suffers from a need for training data annotated on the pixel level (which is tedious and unreliable). We propose a procedure that follows the two-stage model of [16] but requires significantly less training data by virtue of using an unsupervised multi-scale blob detector to generate candidate cell masks. We then select the appropriate masks using a Random Forest classifier trained on a rich feature set and constrained by graph relationships between masks. Our method is:

- spatially adaptive in the face of uneven dye and cell distribution.
- minimally supervised, requiring little training data.
- general between animals, but easily retrained if needed.
- able to segment any cell type, regardless of (in)activity.

2 Analysis Method

Our algorithm takes as input data of the form $F_{d,x,y,t}$ where the value at (d, x, y, t) is a measurement of the fluorescence of dye d at location (x, y) at time t. The goal is to output a set of masks M, where each mask identifies the pixels that sample from a particular cell. In the data used for this paper there are two dyes (one functional showing calcium transients in cells and one structural differentiating types of cells), several thousand time points per recording session, and a spatial resolution of 128×128 pixels covering an area of 240×240 microns.

2.1 Stage 1: candidate mask generation

The first stage of our algorithm is unsupervised and generates a structured set of candidate cell masks. In this preliminary work, we generate these masks from the time-averaged data from the functional dye. Let $\bar{F}_{x,y} = \frac{1}{T}\sum_{t=1}^{T} F_{d^\star,x,y,t}$ be the time averaged image where $d^\star$ is the index of a dye.

Cells are characterized by regions of locally increased fluorescence. We use a Laplacian-of-Gaussian (LoG) multi-scale blob detector to identify these regions. Using a multi-scale approach is critical due to spatial variations in fluorescence intensity and in the size and spacing of cells (see figure 2).

We first smooth the image $\bar{F}_{x,y}$ with a sequence of Gaussian kernels with scales $s \in S$, ranging from no smoothing to over-smoothing. Let $F'_{x,y,s}$ be this set of smoothed images. We then convolve each smoothed image with a 3-by-3 kernel K_L which approximates the Laplacian operator. Since cells are characterized by locally high fluorescence, we select the pixels of the resulting image with negative values as belonging to candidate cell masks. Let the binary images $F^{01}_{x,y,s}$ indicate these pixels.

Finally, we segment the regions selected by the LoG procedure into individual cells. We take the set of local maxima in a smoothed image to be the set of candidate cells c_i at that level of smoothing. For high levels of smoothing there will be very few local maxima, whereas for no smoothing there may be several hundred. We assign each non-zero pixel of F^{01} to the candidate cell (local maxima) that is the termination of a simple hill-climbing process on $F'_{\cdot,\cdot,s}$.

With reasonable step sizes between smoothing levels, it is possible to trace each local maxima through scale space. For this we use a two-dimensional adaptation of the mode tree methodology described in [19]. To summarize, in the maximally smoothed image $F_{\cdot,\cdot,s_i}$ there will be a set of local maxima $c_{1\ldots n}$ where n is typically very small. We assign each of these maxima a unique ID. At the next lower smoothing level, there will be another set of local maxima $c_{1\ldots m}$ where $m \geq n$. Each of the original n maxima can be matched with one of the m maxima, and the appropriate ID transferred. Any remaining maxima from the larger set receive unique IDs. Each new maxima may, in addition, be assigned a parent by hill-climbing from its location in $F_{\cdot,\cdot,s_i}$.

At this point, we have a tree structured collection of candidate cell masks $M_{c,s}$ indexed by cell ID c and smoothing level s. A particular c identifies a set of up to $|S|$ masks, though only a small number of candidate cells will exist at all smoothing levels.

2.2 Stage 2: mask selection

The first stage of our procedure gives us a structured set of candidate cell masks $M_{c,s}$. The majority of these candidate masks should be rejected (in a typical dataset, less than half a percent of the candidate masks are correct). Let $G(M_{c,s})$ be a goodness measure for $M_{c,s}$, indicating how strongly we believe it to be correct. We wish to choose zero or one mask for each candidate cell. Let $L = [l_1, ..., l_C]$ be the vector whose elements are zero or an integer giving the smoothing index of the selected map for each cell c. We aim to select

$$\hat{L} = \arg\max_{L} \sum_{c=1}^{C} G(M_{c,l_c}) \tag{1}$$

subject to constraints that $\hat{L}$ respect the parent-child structure of the set of masks (if a child is selected, the chosen smoothing level for its parent must be below that at which the child originates), and that each selected mask has a sufficiently large goodness value. This is a complex optimization problem, but an approximation that works well in practice is to solve the trivial unconstrained problem and then enforce the goodness threshold and parent-child relationships *post hoc*.

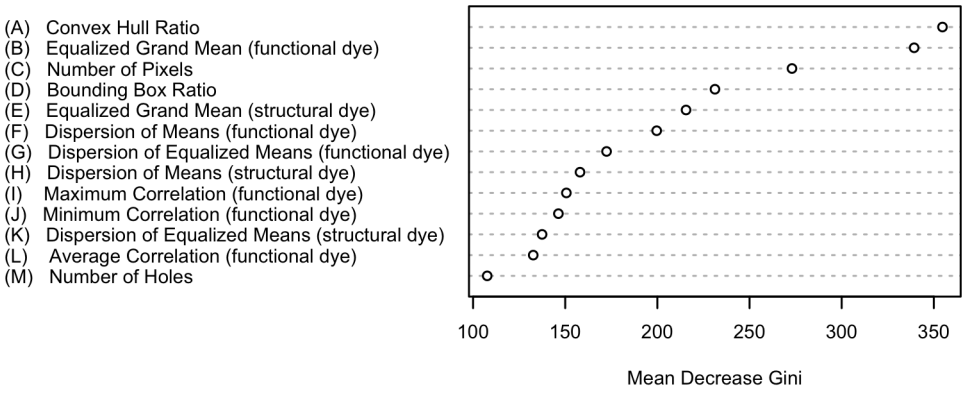

Fig. 1. The set of features used in our Random Forest classifier to calculate the goodness of masks. The plot shows the Mean Decrease Gini (MDG) importance values. Notably, these features include both properties of the mask itself (A,C,D,M), properties of the time-averaged data from multiple dyes (B,E-H,K), and properties of the time series from the functional dye (I,J,L). This is a richer summary of the data than what is typically available to a human annotator.

We use a the estimated class probabilities from a Random Forest classifier [20] to estimate $G(M_{c,s})$. The Random Forest classifier is fast and performs well in the face of extremely unbalanced class frequencies and weakly predictive features without known class conditional distributions. As input to the classifier, we use features that describe both the shape of the mask itself as well as the data covered by the mask. We find that both classes of features contribute significantly to performance. The specific features used in this paper are shown in figure 1, though a strength of our method is that additional features appropriate to the situation can easily be added.

2.3 Application: Rat Somatosensory Cortex

To demonstrate our procedure we use data collected from rat somatosensory cortex. To reduce the effects of brain motion during *in vivo* imaging we align the images in each imaging session using rigid-body translation with correlation based estimates of offset. To remove some of the effect of out-of-plane brain motion (which is primarily driven by respiration and is highly periodic), we apply a high order autoregressive filter to each pixel in the aligned images.

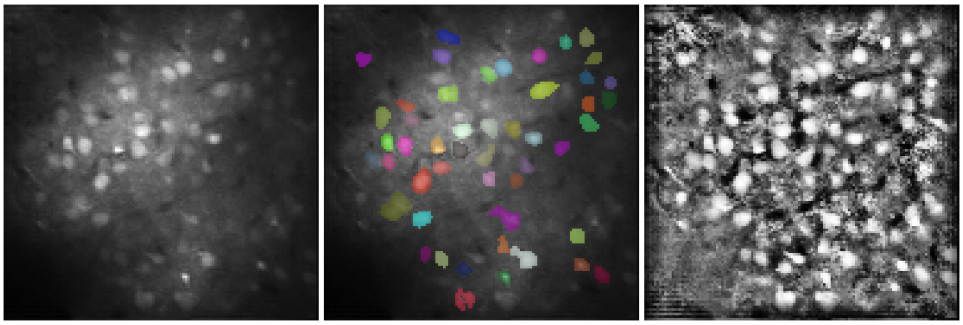

Fig. 2. The output of our completely automated procedure on a dataset from a different animal than those used for training. Left is the time averaged image. Right is the equalized version created using a windowed histogram equalization algorithm. Center is the final results of the algorithm with each mask colored arbitrarily. Desirable properties of this segmentation include that cells are detected both in the bright central region and in the dimmer periphery. In addition, many closely spaced cells are successfully separated. Some errors made in this segmentation include missed cells and some closely spaced cells which are not successfully separated.

For this preliminary work we train our second stage classifier using annotated mask sets from two imaging sessions in one animal (a total of

74 cells). With a custom tool, annotation at the mask level was not time consuming. Figure 2 shows an example of the output of our procedure applied to data from a different animal.

One major challenge of this work is evaluation. There are no standard, carefully hand labeled data sets for comparison. There is no standard method for generating simulated data. We are considering ways to improve this situation. In the meantime, we informally evaluated, on the pixel level, performance on 20 recording sessions from several animals. Our procedure output acceptable masks for about 80% of the cells, with an additional 5% of cells receiving masks with some small problems. Importantly, many of the errors stemmed from failing to generate a correct mask in the first stage of our procedure. This is expected since for this preliminary work we only used one channel and one mask generation method. We expect better performance in future work using all the data available to generate masks.

3 Discussion

We presented here a procedure for automated segmentation of cells in two-photon calcium imaging. The first stage of the procedure uses a multi-scale blob detector to generate a structured set of candidate masks. This is completely unsupervised, and can handle spatially uneven fluorescence intensity and cell distribution. The second stage of our procedure uses a supervised Random Forest to select the appropriate masks. This stage of our procedure performs well by using a rich representation of the mask properties, and the tree structure of the candidate mask set.

There are many improvements to be made to this procedure. We use a rich feature set for our second stage classifier, but we currently use only the time-averaged data from a single dye as input to our blob detector. Using a richer representation of the data throughout the procedure would no doubt improve performance. In addition, there are a practically unlimited number of additional features that could be added. Finally, it is important to evaluate the performance of our procedure in an unbiased, perhaps by comparing it to careful manual segmentations from several experts. Nevertheless, our procedure improves upon existing methods used in the field. It can segment cells regardless of their type or (in)activity, and it requires very minimal training data allowing easy adaptation to different animal types, brain regions, or experimental parameters.

References

1. Lütcke, H., Helmchen, F.: Two-photon imaging and analysis of neural network dynamics. Reports on Progress in Physics **74**(8) (August 2011) 086602
2. Kerr, J.N.D., Denk, W.: Imaging in vivo: watching the brain in action. Nature reviews. Neuroscience **9**(3) (March 2008) 195–205
3. Göbel, W., Helmchen, F.: In vivo calcium imaging of neural network function. Physiology (Bethesda, Md.) **22** (December 2007) 358–65
4. Helmchen, F., Denk, W.: Deep tissue two-photon microscopy. Nature Methods **2**(12) (2005)
5. Yaksi, E., Friedrich, R.W.: Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca 2 + imaging. Nature Methods **3**(5) (2006) 377–383
6. Vogelstein, J.T., Watson, B.O., Packer, A.M., Yuste, R., Jedynak, B., Paninski, L.: Spike inference from calcium imaging using sequential Monte Carlo methods. Biophysical journal **97**(2) (July 2009) 636–55
7. Nimmerjahn, A., Kirchhoff, F., Kerr, J.N.D., Helmchen, F.: Sulforhodamine 101 as a specific marker of astroglia in the neocortex in vivo. Nature Methods **1**(1) (2004) 1–7
8. Lohr, C., Deitmer, J.W.: Calcium Imaging of Glia. In Verkhratsky, A., Petersen, O.H., eds.: Calcium Measurement Methods. Volume 43 of Neuromethods. Humana Press, Totowa, NJ (2010)
9. Reeves, A.M.B., Shigetomi, E., Khakh, B.S.: Bulk loading of calcium indicator dyes to study astrocyte physiology: key limitations and improvements using morphological maps. The Journal of neuroscience : the official journal of the Society for Neuroscience **31**(25) (June 2011) 9353–8
10. Drew, P.J., Shih, A.Y., Kleinfeld, D.: Fluctuating and sensory-induced vasodynamics in rodent cortex extend arteriole capacity. Proceedings of the National Academy of Sciences of the United States of America **108**(20) (May 2011) 8473–8
11. Mishchencko, Y., Vogelstein, J.T., Paninski, L.: A Bayesian approach for inferring neuronal connectivity from calcium fluorescent imaging data. The Annals of Applied Statistics **5**(2B) (June 2011) 1229–1261
12. Dombeck, D.a., Khabbaz, A.N., Collman, F., Adelman, T.L., Tank, D.W.: Imaging large-scale neural activity with cellular resolution in awake, mobile mice. Neuron **56**(1) (October 2007) 43–57
13. Greenberg, D.S., Kerr, J.N.D.: Automated correction of fast motion artifacts for two-photon imaging of awake animals. Journal of neuroscience methods **176**(1) (January 2009) 1–15
14. Bonin, V., Histed, M.H., Yurgenson, S., Reid, R.C.: Local diversity and fine-scale organization of receptive fields in mouse visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience **31**(50) (December 2011) 18506–21
15. Miri, A., Daie, K., Burdine, R.D., Aksay, E., Tank, D.W.: Regression-based identification of behavior-encoding neurons during large-scale optical imaging of neural activity at cellular resolution. Journal of neurophysiology **105**(2) (February 2011) 964–80
16. Valmianski, I., Shih, A.Y., Driscoll, J.D., Matthews, D.W., Freund, Y., Kleinfeld, D.: Automatic identification of fluorescently labeled brain cells for rapid functional imaging. Journal of neurophysiology **104**(3) (September 2010) 1803–11

17. Mukamel, E.A., Nimmerjahn, A., Schnitzer, M.J.: Automated analysis of cellular signals from large-scale calcium imaging data. Neuron **63**(6) (September 2009) 747–60
18. Kwan, A.C.: Toward reconstructing spike trains from large-scale calcium imaging data. HFSP journal **4**(1) (February 2010) 1–5
19. Minnotte, M.C., Scott, D.W.: The Mode Tree : A Tool for Visualization of Nonparametric Density Features. Journal of Computational and Graphical Statistics **2**(1) (1993) 51–68
20. Breiman, L.: Random forests. Machine learning (2001) 5–32