AN ANALYSIS OF TWO WIDELY USED UNSUPERVISED APPROACHES TO SIMULTANEOUSLY EXTRACTING SPATIAL FOOTPRINTS AND FLUORESCENCE TRACES OF INDIVIDUAL NEURONS FROM CALCIUM IMAGING DATA

A PREPRINT

Wenjia Wang
Department of Statistics, University of California, Davis
xxvwang@ucdavis.edu

October 15, 2019

ABSTRACT

This review mainly analyzes two approaches to simultaneously extracting the spatial footprints and fluorescence traces of individual neurons from calcium imaging data: dictionary learning approach (Petersen et al., 2018) and CNMF (Buchanan et al., 2018), and theoretically evaluates them in different scenarios (eg. variation of cell types, field of view, resolution and imaging rate).

1 Introduction

Calcium imaging is an important technique for monitoring the activity of a large number of neurons simultaneously. Repeatedly scanning a single image plane yields a time series of 2D images. This is effectively a video in which neurons blink whenever they are active. In fact, the characteristics of images are different across neuronal compartments recorded (eg. cell body, axonal boutons, dendritic compartments, and spines), field of view (FOV), spatial and temporal resolution. In addition, any method of analyzing calcium imaging data must be tailored to, besides fitting the features of images, the demands of the specific scientific application. For example, some down-stream analysis is to figure out the correlation of spike activities across the populations of neurons, which requires the large-scale recordings.

In recent researches, two classes of methods have received much more attention as tools for analyzing the calcium imaging data: supervised learning approaches (J. Apthorpe et al., 2016; Valmianski et al., 2010) and unsupervised basis learning methods (Mukamel et al., 2009; Maruyama et al., 2014; Pnevmatikakis et al., 2016; Zhou et al., 2016). Independent component analysis (ICA) approach (Mukamel et al., 2009) is one of the first unsupervised methods proposed to handle the calcium imaging data. The well-known drawback is this linear decomposition will fail in the case of significant spatial overlap. PCA/ICA is searching for a matrix of linear spatial filters such that when these filters are applied to the video imaging data, the output time series are as independent as possible, while by contrast the true temporal traces of the significantly overlapping neurons should be highly correlated. Thus the spatial filter tends to exclude many overlapping high-SNR pixels, resulting in noisy and corrupted spiking activity of neurons. In particular, for large-scale imaging data, due to the neuropil activities, the correlation of adjacent neurons will be higher, resulting in failure to separate them by using independent component analysis. The non-negative matrix factorization (NMF) is primative method in image processing (Vogelstein et al., 2010). NMF (Maruyama et al., 2014) outperforms PCA/ICA approach in terms of dealing overlapping neurons, but it does not exploit the special structure of neuronal signals, thus leading to noisy fluorescence estimates.

Currently, dictionary learning approach (Petersen et al., 2018) and CNMF (Buchanan et al., 2018), analyzed in this paper, are two improved and widely-used unsupervised approaches. While they share similar pipeline: denoising (preprocessing), preliminary segmentation (initialization approach) and matrix factorization, the specific methods in each step are different between them. Actually, neither of them is universally effective across different types of imaging data, and a severe limitation of the two approaches is inefficient handling neuropil contamination which is dominant in

large-scale recordings. This paper is organized as introduction to the characteristics of various types of imaging data, analysis of the methodology of the two approaches and discussion on their effectiveness and limitations when applied to different imaging films, and proposal of some feasible improvements.

2 Notation

For mathematical simplicity, each two-dimensional frame in image video is vectorized as a one-dimensional column vector with length P. Let T denote the number of frames of the image video and the raw image data can be expressed as a $P \times T$ matrix $\tilde{Y}$ for which the (x,t) element, $\tilde{Y}(x,t)$, contains the fluorescence of the xth pixel in the tth frame. Let the $P \times P$ matrix M denote the motion correction which transforms the raw data $\tilde{Y}$ to Y^* by $Y^* = M\tilde{Y}$. Considering image pixels can contain a complex mixture of signals from neurons, noise and background, one typical approach is to model data Y^* as $Y^* = AZ + B + E$, where the columns of $A \in \mathbb{R}^{P\times K}$ model the location of each neuron (with K neurons in total), the rows of $Z \in \mathbb{R}^{K\times T}$ model the time-varying fluorescence of each neuron, $B \in \mathbb{R}^{P\times T}$ is a "background" term and $E \in \mathbb{R}^{P\times T}$ denotes noise. Separate Y^* into S patches and denote the s-th patch as Y_s^*. After localized PMD, we obtain $\hat{Y}_s = U_s V_s$, where $\hat{Y}_s$ denotes the denoiser output in the s-th patch, $U_s \in \mathbb{R}^{P\times K_s}$ and $V_s \in \mathbb{R}^{K_s\times T}$ denote the corresponding spatial factor and temporal factor composed of K_s components respectively. Each spatial component in the s-th patch (column vector of U_s) is denoted by u_{s_k} and each temporal component (row vector of V_s) is v_{s_k}. Denote $\hat{Y}$ as the denoised data.

3 Calcium Imaging Data

The quality of calcium images varies over the types of microscopes, calcium indicators and the rate of sampling. Scientists need to carefully choose instrumentation and recording parameters to meet the demands of specific scientific application. For example, if the purpose is to analyze the differential activity such as synaptic inputs to individual spines, integration in dendrites and the probability of release from different boutons, basically sub-cellular resolution is required to detect the activities of individual axonal boutons, dendritic compartments and spines (Pachitariu et al., 2017). For analysis of correlation of neuronal activity across populations, large-field-of-view microscopes with cellular resolution should be used to simultaneously image activities of neurons distributed across multiple brain areas. Consequently, various scientific applications yield different types of imaging data, and accordingly a variety of approaches are designed to extract neuronal signals from those data based on their specific characteristics.

Before evaluating the approaches to extracting neuronal signals from a wide-range of imaging data types, we first analyze the functional characteristics of different imaging data. Figure 1 visually illustrates large-scale recordings with cellular resolution, small-scale dendritic recordings, dendrites and spines image, and neuropil signal. All of them are recorded by two-photon microscopy.

In general, imaging small field of view with cellular resolution and high sampling frequencies is feasible for most microscopes, and subsequently a wide range of approaches have already been proposed and proved efficiently to handle the small-scale imaging data. However, even simple animal behaviors involve multiple brain regions, and thus a large number of wide-field imaging data sets are generated, spurred by the demands of simultaneously imaging activity distributed across multiple brain areas. This subsequently poses a challenge on analyzing the large-scale imaging data. Specifically, scattering and out-of-focus fluorescence which contribute as background rapidly degrade signal and contrast when imaging in multiple brain regions simultaneously (see Figure 1a). In addition, for the imaging data of large and dense populations, another nontrivial source of background is from neuropil fluorescence. Neuronal signals will be contaminated by the neuropil signal from the axons and dendrites of nearby cells (see Figure 1d). Neuropil signal as one of the main sources of the background contamination should be effectively modeled based on a good understanding of its spatial distribution and temporal structure. Visually, neuropil patches are shaped like rings surrounding the somata of adjacent neurons (see Figure 1d), and the decaying time of neuropil after spiking is relatively longer compared to neuron's. In terms of neuropil's functional organization, neuropil activity is strongly correlated to adjacent somatic activity over long distances, and local neuropil patches exhibit more reliable calcium responses compared to adjacent neuronal responses (Lee et al., 2017). Guided by these features, the neuropil spatial footprints can be modeled as a set of spatially-localized basis functions, which exploits the slow variation of each neuropil patch over space.

Study of synaptic transmission and dendritic integration in vivo is also a hot topic in neuronscience. The development of this study relies on acquisition of the activity in neuronal subcellular compartments: dendrites, spines and axonal boutons (See Figure 1c). And in this case, we are not satisfied by extracting somatic signals, but required to recover the fluorescence traces of various neuronal compartments. The small size of subcellular compartments and its different

biophysical characteristics from neuronal somata pose additional challenges on signal extraction from the imaging data, and accordingly those shape-matching based approaches will be out of effect.

4 Methodology

Without losing generality, we assume that the recorded fluorescence in each pixel is a sum of active neuronal signal, background signals originating from global fluorescence, the neuropil or out-of-focus fluorescence, and spatially-spatially independent noise. According to the characteristics of imaging data discussed before, the contamination signals originating in neuropil (denoted by B^f) can be modeled as a weighted sum of spatially localized basis functions: $B^f = \sum_{j=1}^{r} b_j f_j^\top$, where $b_j \in \mathbb{R}^P$, $j = 1, \ldots, r$ are the basis functions, and $f_j \in \mathbb{R}^T$, $j = 1, \ldots, r$ are the corresponding timecourses. Therefore, the fluorescence value after motion correction in the ith pixel and tth frame can be mathematically expressed as:

$$Y^*_{i,t} = \sum_{k=1}^{K} a_k(i) z_k(t) + \sum_{j=1}^{r} b_j(i) f_j(t) + b_0(i) + \epsilon, \tag{1}$$

where a_k is the kth column of spatial matrix A, z_k is the kth row vector of temporal matrix Z, b_0 denotes the spatial component of global background and ϵ is the independent noise (note that the organization of neuropil's spatial component (b_j) and temporal component (f_j) are different from those of neuron (a_j, z_j) respectively).

The typical pipeline of extracting neuronal signals from calcium imaging film consists of motion correction, ROIs identification, recovery of fluorescence traces and optionally extracting spike activities. We assume the data has already been motion corrected. In order to identify the regions of interest more accurately, both the dictionary learning approach (Petersen et al., 2018) and CNMF (Buchanan et al., 2018) involve a denoising process to mitigate the contamination of independent noise. Therefore, both of them can be roughly divided into three main steps: data preprocessing, ROIs identification and signal extraction. In this section, we will analyze and compare the specific methods used in each step by Petersen et al. (2018) and Buchanan et al. (2018).

4.1 Denoising

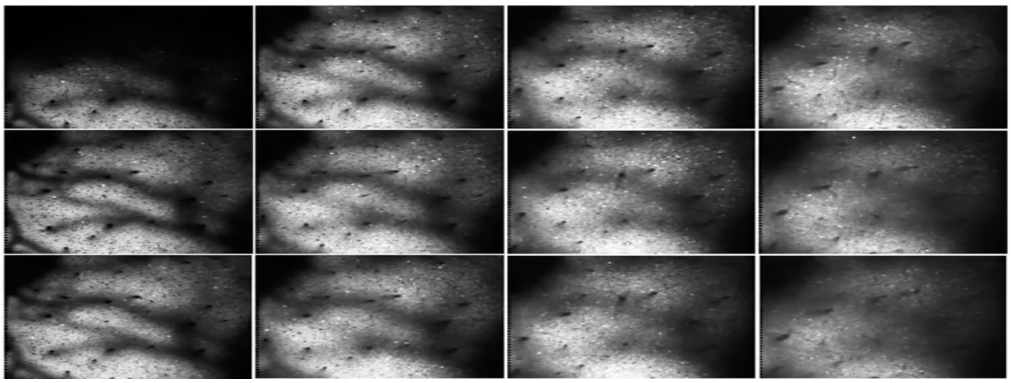
Dictionary learning approach (Petersen et al., 2018) incorporates three steps to preprocess the motion corrected data. First, they smooth the raw $P \times T$ data matrix spatially and temporally using a Gaussian kernel smoother with a bandwidth of one pixel. Second, they adjust for any bleaching effect over time. Specifically, they fit a smoothing spline with 10 degrees of freedom to the median fluorescence for each frame over time, and subtract the frame-specific smoothed median from the corresponding frame. Finally, apply a slight variation of the often-used $\Delta f/f$ transformation, which is for the ith pixel in the tth frame, the standardized fluorescence is equal to

$$y_{it} = \frac{y^*_{it} - M_i}{M_i + q_{10\%}(Y^*)}, \tag{2}$$

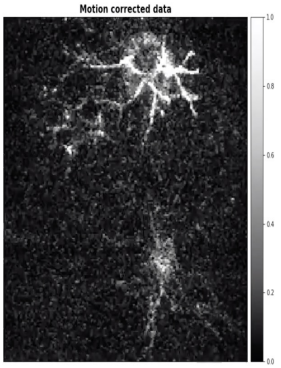
where y^*_{it} is the fluorescence after smoothing and bleaching correction, and M_i is the median of $\{y^*_{ij}\}_{j=1,\ldots,T}$. The author expects that the intensities of "noise pixels" (i.e., pixels that are not part of a firing neuron in that frame) will have a distribution that is approximately symmetric and approximately centered at zero after the preprocessing steps, while non-noise pixels have larger values.

To evaluate the effectiveness of the denoising process, we firstly clarify the results of each step. The Gaussian kernel smoother serves as a low-pass filter to spatially and temporally smooth the image film and reduce the noise. However, a typical shortcoming is that this can make the image film blurry across space and time. Specifically, the dendrites compartments and outlines of neurons tend to be blurred out, and the spike activity may be invisible. The second step and $\Delta f/f$ transformation can remedy bleaching effect over space and time respectively to some extent. Subtracting the frame-specific smoothed median from the corresponding frame can even remove relatively weak global background contamination. And $\Delta f/f$ transformation helps to make the fluorescence at spiking time outstanding. However, these prescriptions can not salvage the dendritic signals (see Figure 1b), and consequently the denoised data only retains the signals of somata. Therefore, we cannot acquire the activity of excitatory synapses distributed across the dendritic arbor (Chen et al., 2013), accordingly hindering the study of synaptic transmission and integration.

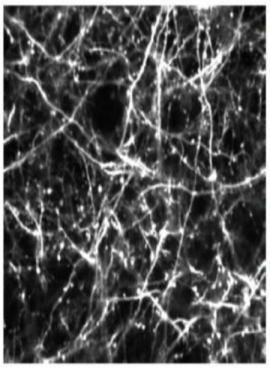
By contrast, the denoising process in CNMF (Buchanan et al., 2018) mainly involves two steps: patch-wise standardizing and penalized matrix decomposition (PMD). The author assumes spatially and temporally uncorrelated noise, then estimates the variance of noise by computing the power spatial density of $y_{i,\cdot}$ followed by averaging its value at the range of high frequencies. This standardization aims to normalize the noise so that we can adaptively choose the matrix rank and extract the leading components which can keep the most neuronal signals.



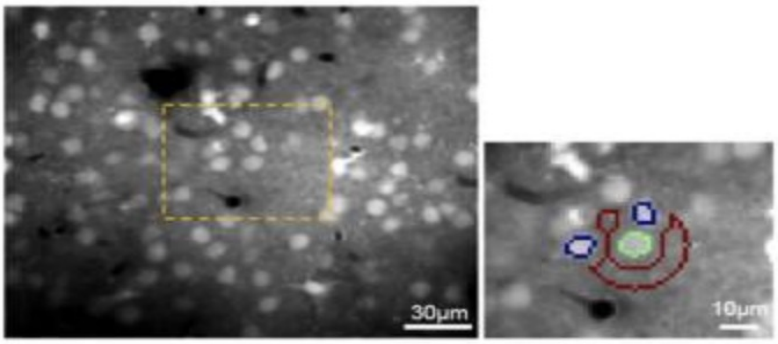
(a) Large-scale imaging data



(b) small-scale imaging data



(c) Dendrites and Spines



(d) Neuropil signals

Figure 1: Visual comparison of various types of imaging data. (a) Imaging of an entire local cortical population, recorded from 11 planes in mouse visual cortex (GCaMP6s expressed in pyramidal cells of all layers using an Ai94; Camk2a-tTA; Emx1-Cre triple transgenic), and the whole video can be seen at https://www.youtube.com/watch?v=xr-flH2Ow2Y. (b) The small-scale two-photon image is from Bessel dendritic imaging data, which not only displays strong somatic components, but also contains visible axonal and dendritic components. Further details can be seen at https://drive.google.com/file/d/1BLx2aEBkKCJDO7x-mB6iIY5kLcvrhvH8/view. (c) Average projections of motion corrected data from cortical layer 1 acquired at high magnification, at 20m vertical spacing. Dendrites were labelled with GCaMP6s and tdTomato (Pachitariu et al., 2017). (d) Selecting cell somata and local neuropil patches. The orange area is magnified on the right. In the right example, the green outline shows a selected cell soma. The corresponding local neuropil patch was defined around the cell body as an annulus (outlined in red). Note that nearby neurons were excluded from the neuropil annulus (outlined in blue)(Lee et al., 2017).

Compared to the preprocessing method involved in dictionary learning approach, patch-wise PMD is more conservative at the cost of computational complexity. Standardizing the noise can guarantee that temporally and spatially correlated signals (including neuronal activities and background contamination) account for the most variance, and thus implementing penalized matrix decomposition will keep all neural signals and leave the majority of noise in residuals. Actually, this is relatively conservative, since background contamination and weak noise may simultaneously remain when keeping the complete neuronal signals. In addition, the local PMD method may be unfeasible for large-scale data containing dense neurons. Even parallel computation is accessible for patch-wise PMD, each patch may still contain a large number of neurons, which implies high rank. Thus, adaptively choosing the matrix rank in penalized matrix decomposition may consume much more time.

4.2 Preliminary Segmentation

Preliminary segmentation (ROIs identification) also serves as the initialization for the following matrix factorization step. The dictionary learning approach firstly identifies the spatially connected components by using the notion of 4-connectivity in each frame, and then clusters these components based on their spatial and temporal dissimilarity. By comparison, Buchanan et al. (2018) identifies the super pixels by calculating the pair-wise correlation of temporal fluorescence of neighboring pixels, and then filters the pure pixels by using the Successive Projection Algorithm (SPA).

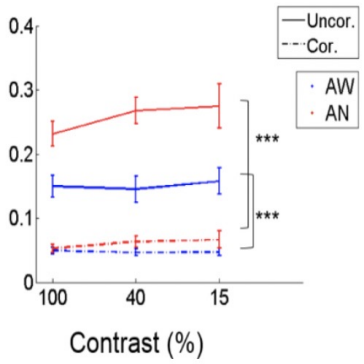
Theoretically, the segmentation approach in Buchanan et al. (2018) outperforms the first paper in terms of the ability to deal overlapping neurons, since the SPA is able to discard those mis-identified overlapping neurons. In addition, the dictionary learning approach actually can instead directly exploit the spatial and temporal information to identify components. There is no need to preliminary detect the spatially connect components from each frame separately and then cluster them over frames. We can instead use the temporal correlation of some neighboring pixels to identify the neuron candidates. Notably, the neighboring pixels should be carefully defined. One common drawback of the two approaches is that the segmentation is based on the assumption that individual neurons are locally localized. However, the dendrites compartments belonging to one neuron are not locally distributed (see Figure 1c). Actually it's good to detect multiple connected components separately. If they are from the same neuron, their temporal correlations must be high (e.g., larger than 0.8), then we can merge them later. However, this may be infeasible, because that in preprocessing step, those dendrites compartments may be already filtered away.

Another common drawback is that both of the two methods ignore the background contamination originating from the local perisomatic neuropil signals due to low z-resolution. This may result in failure to deconvolve neuronal signals since neuropil's signal can notably dominate the temporal correlation between nearby neurons. Consequently, the correlation-based ROIs are not reliable without removing the neuropil contamination. Considering the functional organization of neuropil signal in large-scale data we discussed before, because the fluorescence value at each pixel is dominated by neuropil signal which spatially decays over long distance and is highly correlated over field of view, the correlation of temporal fluorescence between nearby neurons tends to be high (see Figure 2a), and thus they are more likely to be identified as one neuron. In addition, due to the high correlation of neuronal signal to neighboring neuropil-patch signal, the final estimate of fluorescence trace may be the mixed signals from the neuron and neuropil patch. As we discussed before, the decaying time of neuropil responses tends to be longer than neuronal signal, and thus some neuronal spiking activities may be submerged under noisy fluorescence (see Figure 2c). Consequently, the down-stream inference of neuronal spiking activity becomes challenging. However, only high levels of neuropil contamination begin to distort visually evoked responses (Lee et al., 2017), and empirically such high levels of contamination are not common (see Figure 2b). Therefore, the neuropil contamination mainly reflects on dominating the correlation of pairs of pixels in practice, subsequently leading to mis-identification of ROIs.

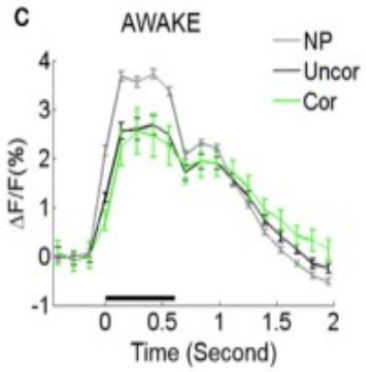
In fact, before identifying the super pixels, Buchanan et al. (2018) soft-thresholds the denoised data, which serves to extract most spiking activity from functional imaging data. This benefits reduction of the neuropil contamination to some extent. Specifically, assume z_i and z_k are fluorescence traces of two adjacent neurons, and f_j denotes their neighboring neuropil-patch fluorescence trace. According to the functional organization of neuropil's signal, f_j is highly correlated to z_i and z_j simultaneously, but these high correlations are mainly contributed by the long decaying property of neuropil signal but not most spiking activities. Thus, correlation based on the extracted most spiking activity is helpful to distinguish different neurons. However, the effectiveness of this remedy needs to be proved in future study.

4.3 Matrix Factorization

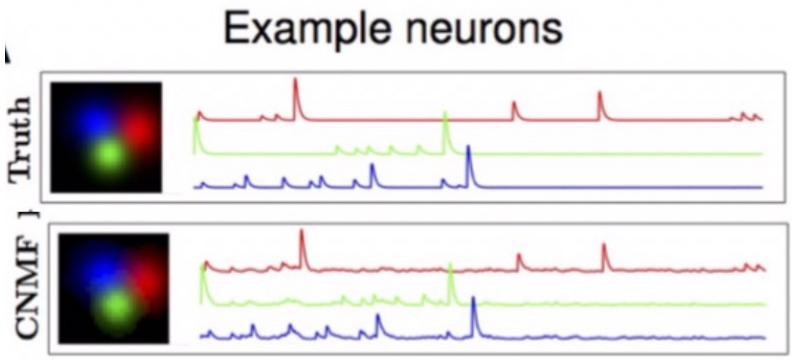
Generally, matrix factorization is a non-convex problem used to estimate the spatial and temporal signals simultaneously. However, dictionary learning approach estimates the temporal signals associated with the K^f elements in the final dictionary $A^f \in \mathbb{R}^{P\times K^f}$ by fixing the spatial matrix. In this way, the problem alters to be a convex regressing



(a) Correlation dominated by neuropil



(b) Response of one cell and neuropil patch



(c) Noisy trace corrupted by neuropil signal

Figure 2: Illustrate the effect of neuropil contamination on fluorescence estimates : (a) Comparison of correlation of pairs of neurons before and after neuropil contamination correction. “UnC." and “C." represent no correction of neuropil contamination and correction of neuropil contamination respectively. *** indicates a significance level of 1e-5 (Lee et al., 2017). (b) Example of the average response of cells and neuropil-patches in a single FOV. “NP”: represent neuropil response. “Uncor”: represent somatic response without neuropil contamination correction. “Cor”: represent somatic response after neuropil contamination correction (Lee et al., 2017). (c) The upper row is the true spiking activities of three neighboring example neurons. The lower row is the fluorescence estimates by using CNMF proposed by (Pnevmatikakis et al., 2016) , which recovers temporal traces that are sufficiently noisy that small calcium transients are submerged in the noise (Zhou et al., 2016).

optimization with non-negative constraint, lasso and grouped lasso penalties, which is much cheaper:

$$\min_{Z\in\mathbb{R}^{K^f\times T}, Z\geq 0}\{\frac{1}{2}\|Y-\tilde{A^f}Z\|_F^2+\lambda\alpha\sum_{k=1}^{K^f}\|z_{k,.}\|_1+\lambda(1-\alpha)\sum_{k=1}^{K^f}\|z_{k,.}\|_2\}, \tag{3}$$

where $\tilde{A^f}$ consists of the scaled $a^f_{.,i}$, which means $\tilde{a^f_{.,i}}=a^f_{.,i}\,\|a^f_{.,i}\|_2^2$.

This essentially solves the near-separable NMF problem based on sparse regression and self dictionaries(Gillis and Luce, 2018). In fact, except involving another lasso penalty term, this is almost equivalent to the successive projection algorithm (SPA) used for initialization by Buchanan et al. (2018).

$$\min_{Z\in\mathbb{R}_+^{K^f,T}}\|Z\|_{1,2}\ \text{s.t.}\|Y-\tilde{A^f}Z\|_F^2\leq\epsilon\ \text{and}Z\leq 1, \tag{4}$$

where $\|Z\|_{1,1}=\sum_{i=1}^{K^f}\|z_{i,.}\|_2$. Additionally, the regression step of dictionary learning approach incorporates another constraint on the sparseness of row vectors of Z. Thus the method of computing “pure” superpixels (Buchanan et al., 2018) and the method of refining neural dictionary elements (Petersen et al., 2018) are closely related. However, this regression step involves tuning parameter λ by train-validation set among each connective patch as the author suggested. In each patch, the number of pixels is much smaller, tuning parameter by dividing the pixels into training and validation sets may not be convincing. CNMF does not solve non-convex problems directly, instead it divides the problem into

multiple convex subproblems. Intuitively, dictionary learning may be faster than CNMF, but this needs to be verified by simulation experiments.

However, both the two approaches are unable to correct complicated neuropil contamination in final matrix factorization. When we decompose the spatial and temporal imaging matrix without special constraints on temporal components, highly correlated fluorescence traces from different neurons due to neuropil contamination are likely to be estimated as one component, resulting in failure to deconvolve neuronal signals. Besides, because the neuronal signal is strongly correlated to its neighboring neuropil-patch signal, the recovered fluorescence trace is sufficiently noisy, and subsequently distorts the conclusion in down-stream analysis (details are explained in "Preliminary Segmentation" section).

5 Improvements and Future Study

Based on the above analysis, both the dictionary learning approach and the CNMF are not effective when the imaging data contaminated by complex background. The rough idea to handle neuropil contamination is to remove the neuropil signal or to explicitly recover it. In order to reduce the neuropil contamination in preprocessing data, we can exploit the spatial organization of neuropil signal. For example, the neuropil-patch signal surrounds the neuron and decays slowly over space. Guided by the relatively higher spatial coherence, we can attempt high-pass filter to reduce neuropil contamination (Zhou et al., 2016). From another perspective, if we retain the neuropil signal in the fluorescence matrix and plan to extract the neuronal signals by factorizing the matrix, some constraints on neuronal signal and neuropil signal should be involved to recover them separately. For example, guided by the assumption that the decaying time of neuropil after spiking is longer than neuronal signals, we can try autogressive model to recover neuronal signals by set the decaying parameter p smaller (Vogelstein et al., 2010). However, even Lee et al. (2017) proposed some functional structures of neuropil signal, the characteristics of neuropil activity empirically vary across the state of animals, calcium indicators, microscopes and recording parameters (eg. smapling rate). Hence, the existing models (Zhou et al., 2016; Pachitariu et al., 2017) to present the neuropil signal may be inappropriate in some specific cases. Importantly, study of the activity of populations from multiple brain regions requires the advert of advanced methods to efficiently deal the large-scale imaging data which is more likely to come with complex background due to low z-resolution, so a further study of modeling neuropil signal based on a better understanding its functional structure is worthwhile. Another bottleneck in large-scale imaging film processing is the high cost of computation. For dense population, the number of components in a small patch is still large. Thus, study of speed-up algorithms (Friedrich et al., 2017a,b) is also crucial to analyze the large-scale data.

Spurred by the demands of studying synaptic transmission and dendritic integration in vivo, calcium imaging should not be restricted to somata, but can also record the activity in neuronal subcellular compartments: dendrites, spines and axonal boutons. Biophysically, thousands of synapses distribute across the dendritic arbor. When synapses excite, calcium will accumulate in individual dendritic spines (Chen et al., 2013). To study the synapic transmission and dendritic integration, we need to extract the activity in dendritic spines. This is limited to the imaging technique and the experimental design, it requires expressing calcium indicators only in the spine and fast imaging at high-spatial resolution (300nm). Otherwise, the time courses of separated spines belonging to the same parent dendrite or axon will be highly correlated. In this case, novel statistical methods may be required. Additionally, as we discussed before, dictionary learning approach may filter away those dendrites compartments in preprocessing step, and consequently is unable to identify them. In the future, we can study the methods of efficiently extracting dendrites. Besides, we should also check the effectiveness of existing methods when applied to extract the activity in dendritic spines, and improve them if needed.

References

E. Kelly Buchanan, Ian Kinsella, Ding Zhou, Rong Zhu, Pengcheng Zhou, Felipe Gerhard, John Ferrante, Ying Ma, Sharon H. Kim, Mohammed A Shaik, Yajie Liang, Rongwen Lu, Jacob Reimer, Paul G Fahey, Taliah N Muhammad, Graham Dempsey, Elizabeth Hillman, Na Ji, Andreas S Tolias, and Liam Paninski. Penalized matrix decomposition for denoising, compression, and improved demixing of functional imaging data. *bioRxiv*, 2018. doi: 10.1101/334706. URL https://www.biorxiv.org/content/early/2018/06/03/334706.

Tsai-Wen Chen, Trevor J Wardill, Yi Sun, Stefan Robert Pulver, Sabine Renninger, Amy Baohan, Eric R Schreiter, Rex A. Kerr, Michael B. Orger, Vivek Jayaraman, Loren L Looger, Karel Svoboda, and Douglas S Kim. Ultra-sensitive fluorescent proteins for imaging neuronal activity. 2013.

Johannes Friedrich, Weijian Yang, Daniel Soudry, Yu Mu, Misha B. Ahrens, Rafael Yuste, Darcy S. Peterka, and Liam Paninski. Multi-scale approaches for high-speed imaging and analysis of large neural populations. *PLOS Computational Biology*, 13(8):1–24, 08 2017a. doi: 10.1371/journal.pcbi.1005685. URL https://doi.org/10.1371/journal.pcbi.1005685.

Johannes Friedrich, Pengcheng Zhou, and Liam Paninski. Fast online deconvolution of calcium imaging data. *PLOS Computational Biology*, 13(3):1–26, 03 2017b. doi: 10.1371/journal.pcbi.1005423. URL https://doi.org/10.1371/journal.pcbi.1005423.

N. Gillis and R. Luce. A fast gradient method for nonnegative sparse regression with self-dictionary. *IEEE Transactions on Image Processing*, 27(1):24–37, Jan 2018. ISSN 1057-7149. doi: 10.1109/TIP.2017.2753400.

Noah J. Apthorpe, Alexander J. Riordan, Rob E. Aguilar, Jan Homann, Yi Gu, David W. Tank, and Hyunjune Seung. Automatic neuron detection in calcium imaging data using convolutional networks. 06 2016.

Sangkyun Lee, Jochen F. Meyer, Jiyoung Park, and Stelios M. Smirnakis. Visually driven neuropil activity and information encoding in mouse primary visual cortex. *Frontiers in Neural Circuits*, 11:50, 2017. ISSN 1662-5110. doi: 10.3389/fncir.2017.00050. URL https://www.frontiersin.org/article/10.3389/fncir.2017.00050.

Ryuichi Maruyama, Kazuma Maeda, Hajime Moroda, Ichiro Kato, Masashi Inoue, Hiroyoshi Miyakawa, and Toru Aonishi. Detecting cells using non-negative matrix factorization on calcium imaging data. *Neural Networks*, 55:11 – 19, 2014. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2014.03.007. URL "http://www.sciencedirect.com/science/article/pii/S0893608014000707".

Eran A. Mukamel, Axel Nimmerjahn, and Mark J. Schnitzer. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron*, 63(6):747 – 760, 2009. ISSN 0896-6273. doi: https://doi.org/10.1016/j.neuron.2009.08.009. URL "http://www.sciencedirect.com/science/article/pii/S0896627309006199".

Marius Pachitariu, Carsen Stringer, Mario Dipoppa, Sylvia Schröder, L. Federico Rossi, Henry Dalgleish, Matteo Carandini, and Kenneth D. Harris. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv*, 2017. doi: 10.1101/061507. URL https://www.biorxiv.org/content/early/2017/07/20/061507.

Ashley Petersen, Noah Simon, and Daniela Witten. Scalpel: Extracting neurons from calcium imaging data. *Ann. Appl. Stat.*, 12(4):2430–2456, 12 2018. doi: 10.1214/18-AOAS1159. URL "https://doi.org/10.1214/18-AOAS1159".

EftychiosÂ A. Pnevmatikakis, Daniel Soudry, Yuanjun Gao, Timothy A. Machado, Josh Merel, David Pfau, Thomas Reardon, Yu Mu, Clay Lacefield, Weijian Yang, Misha Ahrens, Randy Bruno, Thomas M. Jessell, DarcyÂ S. Peterka, Rafael Yuste, and Liam Paninski. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron*, 89(2):285 – 299, 2016. ISSN 0896-6273. doi: https://doi.org/10.1016/j.neuron.2015.11.037.

Ilya Valmianski, Andy Y. Shih, Jonathan D. Driscoll, David W. Matthews, Yoav Freund, and David Kleinfeld. Automatic identification of fluorescently labeled brain cells for rapid functional imaging. *Journal of Neurophysiology*, 104 (3):1803–1811, 2010. doi: 10.1152/jn.00484.2010. URL https://doi.org/10.1152/jn.00484.2010. PMID: 20610792.

Joshua T. Vogelstein, Adam M. Packer, Timothy A. Machado, Tanya Sippy, Baktash Babadi, Rafael Yuste, and Liam Paninski. Fast nonnegative deconvolution for spike train inference from population calcium imaging. *Journal of Neurophysiology*, 104(6):3691–3704, 2010. doi: 10.1152/jn.01073.2009. URL https://doi.org/10.1152/jn.01073.2009. PMID: 20554834.

Pengcheng Zhou, Shanna L. Resendez, Jose Rodriguez-Romaguera, Jessica C. Jimenez, Shay Q. Neufeld, Garret D. Stuber, Rene Hen, Mazen A. Kheirbek, Bernardo L. Sabatini, and Robert E. Kass. Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *arXiv e-prints*, art. arXiv:1605.07266, May 2016. URL https://ui.adsabs.harvard.edu/abs/2016arXiv160507266Z.