CRCNS.org alm-2 data description

Version 0.6 (Jan 11, 2016)

Summary

These dataset contains calcium imaging data from the anterior motor cortex (ALM) in mice performing a tactile delay-response task. The raw data consists of images, while the processed data consists of fluorescence transients of individual ROI, the identity of imaged cells (PT or IT), and behavioral variables (object location, licking direction etc.). The data is from four animals and includes 59 recording sessions. Results from the experiments are described in:

> A motor cortex circuit for motor planning and movement
> Nuo Li, Tsai-Wen Chen, Zengcai V. Guo, Charles R. Gerfen & Karel Svoboda
> *Nature* 519, 51–56 (05 March 2015) doi: 10.1038/nature14178

Animals and surgery

To label intratelencephalic cells for imaging, cholera toxin subunit B (CTB; Alexa 647; Molecular probe, Invitrogen, 0.5% in HEPES buffered saline) was injected to the contralateral (right) ALM (2.5 mm anterior, 1.5 mm lateral to bregma, 300 and 600 µm deep, 100 nl per site). For pyramidal tract cells, red RetroBeads (Lumafluor) were injected into the ipsilateral (left) basal pontine nucleus (3.5 mm posterior, 0.4 mm lateral, 5, 5.4, and 5.8 mm below brain surface, 100 nl per site). Window surgery and GCaMP virus injections were carried out 12–34 days after tracer injection. A circular craniotomy (~3 mm diameter) was made above left ALM (centred at 2.5 mm anterior and 1.5 mm lateral to bregma). AAV2/1-syn-GCaMP6s-WPRE virus (UPenn Viral Core, AV-1-PV2824) was diluted two- to sixfold in HEPES buffered saline. Injections were made at three to five locations centred around ALM (separated by ~400 µm) and at three depths (210/370/530 µm) for each location (~5–6 nl per depth). The imaging window was constructed from two layers of microscope coverglass and fixed to the skull using cyanoacrylate glue and dental acrylic. A metal post for head fixation was implanted posterior to the window using dental acrylic. Water restriction started 5–7 days after window surgery. Behavioural training started ~5–7 days after water restriction.

Imaging experiments started after the animals had learned the task (>70% trials correct; typically ~20–30 days after surgery). Images were acquired using a custom-built two-photon microscope equipped with a resonant galvo scanning module (Thorlabs), controlled by ScanImage 4.2 (http://www.scanimage.org). The light source was a femtosecond pulsed laser (Coherent). The objective was a 16× water immersion lens (Nikon, 0.8 NA, 3 mm working distance). GCaMP6s was excited at 940 nm and images (512 × 512 pixels, 400 mm × 400 mm or 600 mm × 600 mm) were acquired at 15 Hz. The average excitation power was up to 120 mW for L5 neurons. After functional imaging of a particular *z* plane, the laser wavelength was switched to 830 nm to image Retrobead and CTB-647. A small image stack was acquired around the imaging location to allow unambiguous identification of cells that were out of focus. Retrobead and CTB-647 were imaged with Brightline 609/54 (Semrock) and HQ-675/70-2P (Chroma), respectively.

Experimental method

Images were acquired using a custom-built two-photon microscope (openwiki.janelia.org/wiki/display/shareddesigns/MIMMS) equipped with a resonant galvo scanning module (Thorlabs), controlled by ScanImage 4.2 (http://www.scanimage.org). The light source was a femtosecond pulsed laser (Coherent). The objective was a 16× water immersion lens (Nikon, 0.8 NA,

3 mm working distance). GCaMP6s was excited at 940 nm and images (512 × 512 pixels, 400 µm × 400 µm or 600 µm × 600 µm) were acquired at 15 Hz. The average excitation power was up to 120 mW for L5 neurons. After functional imaging of a particular *z* plane, the laser wavelength was switched to 830 nm to image Retrobead and CTB-647. A small image stack was acquired around the imaging location to allow unambiguous identification of cells that were out of focus. Retrobead and CTB-647 were imaged with Brightline 609/54 (Semrock) and HQ-675/70-2P (Chroma), respectively.

Data analysis

The brain motion was corrected using a line-by-line correction algorithm. Regions of interest (ROIs) corresponding to identifiable cell bodies were selected using a semi-automated algorithm. Individual neurons were visually identified on average fluorescence images as well as a pixel-based response map and a 'neighborhood correlation map' (where the brightness of each pixel encodes the correlation of its fluorescent time course to that of its immediate neighbours) that highlight task-related and active cells, respectively. The fluorescence time course of each cell was measured by averaging all pixels within the ROI, with a correction for neuropil contamination. The fluorescence signal of a cell body was estimated as $F_{cell\_true}(t) = F_{cell\_measured}(t) - r \times F_{neuropil}(t)$, with $r = 0.7$. The neuropil signal $F_{neuropil}(t)$ surrounding each cell was measured by averaging the signal of all pixels within a ~20 µm region from the cell centre (excluding all selected cells). To ensure robust neuropil subtraction, only cells that were at least 5% brighter than the surrounding neuropil were included. $\Delta F/F_0$ was calculated as $(F–F_0)/F_0$, where F_0 is the baseline fluorescence signal averaged over a 0.5 s period immediately before the start of each trial.

Task-related neurons were defined as neurons showing significant fluorescence modulation during the task. This was calculated using non-parametric ANOVA (Kruskal–Wallis test) across multiple 0.33 s time bins (five image samples) during the task. A cell was classified as task-related if the null hypothesis that all time bins had equal fluorescence can be rejected at a P value of 0.01 during either contra or ipsi trials. This criterion identified 58% (2,740/4,706) cells as task related. We further defined trial-type-selective cells as a subset of task related neurons that showed significantly different $\Delta F/F$ response during contra and ipsi trials ($P < 0.05$; Wilcoxon rank sum test). 57% (1575/2740) of task related cell are trial-type selective.

Data file organization (what's included):

- Processed data files are in the folder "datafiles" (in archive file "datafiles.tar.gz").

Each file "data_XXX.mat" contains the data from one experimental session (i.e. one behavioral session); (e.g. data_an019_2013_08_15_480.mat contains data from animal #19 collected on 2013/08/15 at depth 480um from cortical surface)

Each file "meta_XXX.mat" contains the meta data information for each session.

There is one "meta_XXX.mat" file for each "data_XXX.mat" file.

- A collection of analyses scripts that reproduces Figure 4 in "Li, Chen, Guo, Gerfen, Svoboda, Nature (2015)" is included in the folder "analyses_scripts". To run a script see "How to get started" section below.

- Raw images of each session are located in the folder "raw_images"

Data format

Meta Data File (".mat")

See general description from the lab.

Processed Object File (".mat")

Each .mat data file contains data from one session. The data is in the format of matlab structure. Each structure contains the following fields:

Top level data description

timeUnitIds: This is an vector of integers, with the following convention: 1--ms; 2--second; 3--minute; 4--hour; 5—day
timeUnitNames: Description of time units in **timeUnitIds** (e.g. "second")

Behavior data

trialTimeUnit: what time unit is the data in (refers to **timeUnitIds** above).
trialTypeStr: description of the rows in **trialTypeMat** (e.g. "HitL","ErrL"); by convention, the *correct* response is specified in the directional term (L or R). Thus, error lick left actually produced a lick on the right lickport, but the correct response would have been to lick left. Trials can either be hit (correct), err (incorrect), or nolick (animal made no response).
trialTypeMat: each column describes one trial by the description in **trialTypeMat** (e.g. a "lick left" trial in which the animal correctly reported choice will have a entry of "1" for "Correct lick left" and "0" for "Error lick left". A phostostimulation trial will have "1" for "Photostimulation").
can't find photostimulation
trialIds: trial number to reference to the trials
trialStartTimes: start time of the trial. The time is referenced to session start, (i.e. time 0 is the start of the session).
trialPropertiesHash: contains detailed information about trial structures and timing information. It has the following sub fields:
trialPropertiesHash.keyNames: {'PoleInTime' 'PoleOutTime' 'CueTime' }
trialPropertiesHash.descr: describes entries in **keyNames**
trialPropertiesHash.value: contains the values of the properties in **keyNames** for each trial.

Imaging data

timeSeriesArrayHash: this structure contains data that are time series. The times are relative session start.
timeSeriesArrayHash.keyNames: name of data stored in each hash, {'fmean_roi', 'fmean_neuropil'} represent average fluorescence of individual ROI and surrounding neuropil region, respectively.
timeSeriesArrayHash.descr: more detailed description of what is in each hash

timeSeriesArrayHash.value: the actual data

Each **timeSeriesArrayHash.value** has the following fields (this is a single time series array):

ids: a numerical unique identifier for each data vector
idStrs: a description of what that particular vector is
idStrsDetailed: even more detailed description if needed
valueMatrix: the data ; n x t matrix, where n corresponds to id, idStr, etc., while t corresponds to the time dimesion.
timeUnit: the time unit used for this data
time: time basis for this data, same size as second dimension of valueMatrix, and relative to first trial being 0
trial: same length as time, tells you the trial to which every time point belongs to (for rapid referencing)
depth: the depth of imaging plane (in micrometers)
cellType: the identity of cell 'p' for PT, 'I' for IT, '' for unidentified.
Pixel_list : index of pixels belonging to each ROI

Raw image files
Calcium image files (.tif format). These are stored in the "raw_images" subdirectory, and packaged into .tar.gz files. The file names have the format "anxxx_year_month_date_main_yyyum_RGB.tiff", where xxx indicates the animal number and yyy indicates the imaging plane depth. Imaging plane 1 is the fly back frame. The other image planes (2, 3 and 4) contain Regions Of Interests (ROIs) that are used to extract time series data that are stored in the .mat files.

How to get started

We have included the following demo files to get started with this dataset in 'analysis_scripts' folder. To run a script, (e.g. "Demo_example_fovs.m"):

1) open MATLAB
2) change current folder to the ".\analysis_scripts\" folder
3) type run('Demo_example_fovs.m')

Demo_example_fovs.m– This file contains codes to read cell locations and GCaMP/tracer images from a particular data file

Demo_figure_example_ROIs.m - This file contains analysis codes to reproduce figure 4e of Li, Chen et. al. 2015 (Nature) paper, showing calcium response traces of several example ROI in mice anterior-lateral motor cortex (ALM) during a left-right delay response task

Demo_figure_summary.m - This file contains analysis codes to reproduce figure 4 g,h of Li, Chen et. al. 2015 (Nature) paper, showning calcium signals of contra and ipsi selective neurons among PT and IT subpopulation in ALM

How to cite the data

If you publish any work using the data, please cite the Nuo et al. (2015), publication above and also cite the data set in the following recommended format:

Tsai-Wen Chen, Nuo Li, Charles R Gerfen, Zengcai V. Guo, Karel Svoboda (2016); Calcium imaging responses from anterior lateral motor cortex (ALM) neurons of adult mice performing a tactile decision behavior. CRCNS.org. http://dx.doi.org/10.6080/K04M92GX