Calcium transient detection harnessing spatial similarity (CATHARSiS)

Reference

Jürgen Graf*, Vahid Rahmati*, Myrtill Majoros, Otto W. Witte, Christian Geis, Stefan J. Kiebel, Knut Holthoff#, Knut Kirmse#. (2021) Network instability dynamics drive a transient bursting period in the developing hippocampus in vivo. *bioRxiv*. doi.org/10.1101/2021.05.28.446133.
| * equal contribution | # senior authors
Corresponding author: knut.kirmse@uni-wuerzburg.de

Objective

CATHARSiS was devised for detecting somatic Ca^{2+} transients (CaTs) in GCaMP6s-expressing neurons in the developing CA1 *in vivo*, where optical overlap of neighboring cells can constitute a major source of false positive events. By making use of both the amplitude and spatial distribution of fluorescence changes (ΔF_{xy}), CATHARSiS can achieve a higher precision than methods relying on mean (i.e. spatially averaged) ΔF alone. It is this "cleansing" from spurious events that has given the method its name (catharsis = κάθαρσις).

Method

For details of the method, we refer the reader to the above manuscript. The method consists of three major steps: (1) constructing a spatial ΔF_{xy} template representing the active cell; (2) computing a detection criterion D(t) by comparing this spatial ΔF_{xy} template to the measured ΔF_{xy} in each frame; (3) extracting CaT onsets from D(t).

Ad (1): For each region of interest (ROI), we first obtain the mean F(t) by frame-wise averaging across all pixels of that ROI. We then compute the first derivative of F(t) and smooth it using a Savitzky-Golay algorithm, yielding $\dot{F}(t)$. We next determine a number N of candidate CaT onsets by extracting the frame numbers corresponding to the N $\dot{F}(t)$ peaks having the largest amplitude. This step is performed in an iterative-descending manner by starting with the largest $\dot{F}(t)$ peak. For each peak, we define a minimum time difference to all subsequently extracted peaks, so as to avoid extracting nearby frames belonging to the same CaT. For each candidate CaT onset, we then compute the corresponding spatial ΔF_{xy}. To this end, optionally, raw ROIs can be radially expanded using a Euclidian distance transform (this might increase detection reliability by including the contrast-rich transition from cytoplasm to surrounding neuropil). Resting fluorescence $F_{0,xy}(t)$ is obtained by oversmoothing $F_{xy}(t)$. The N candidate ΔF_{xy} templates are computed by converting raw ΔF_{xy} values into z-scores. Based on visual inspection in a GUI, the user is expected to deselect those candidate ΔF_{xy} templates that putatively reflect activation of optically overlapping somata and/or neurites. A given cell is included in the further analysis, if a user-defined minimum number of candidate ΔF_{xy} templates is selected; otherwise, the cell is excluded from

analysis. For each included cell, the selected candidate ΔF_{xy} templates are averaged to obtain the final ΔF_{xy} template representing the active cell.

Ad (2): For each ROI (optionally expanded as described above), we extract its spatial ΔF_{xy} for all frames in the image stack. Next, the spatial ΔF_{xy} template representing the active cell is optimally scaled to fit its ΔF_{xy} in each recorded frame. Based on the optimum scaling factor and the goodness of the fit, a detection criterion D(t) is computed for each time point. D(t) was defined without modification as previously described for a template-matching approach in the temporal domain (Clements and Bekkers, 1997).

Important

CATHARSiS makes use of the fact that, while a given cell is active, ΔF_{xy} is spatially inhomogeneous ('ring-like' for a typical GCaMP-expressing cell soma) and characteristic of this cell. Therefore, in simplifying terms, a high-contrast ΔF_{xy} template is a good template. By design, D(t) is unaffected by spatially homogeneous ΔF_{xy}, e.g. due to diffuse changes in background fluorescence.

Ad (3): For each ROI, CaT onsets are extracted from D(t) using UFARSA, a general-purpose event detection routine (Rahmati et al., 2018). To this end, we slightly modified the original UFARSA approach in two ways. I) Following the smoothing step implemented in UFARSA, all negative values are set to zero, as negative-to-positive transitions occasionally result in false positive events. II) We introduce a lower bound for the leading threshold, so as to minimze potential false positive events. Reconstructed CaT onsets are translated into a binary activity vector.

References

1. Clements JD, Bekkers JM (1997) Detection of spontaneous synaptic events with an optimally scaled template. Biophys J 73:220-229.
2. Rahmati V, Kirmse K, Holthoff K, Kiebel SJ (2018) Ultra-fast accurate reconstruction of spiking activity from calcium imaging data. J Neurophysiol 119:1863-1878.

Software required

(1) Matlab (MathWorks), tested with Matlab R2021b

(2) ImageJ/Fiji, tested with Fiji 1.53c and Java version "1.8.0_333", Java(TM) SE Runtime Environment (build 1.8.0_333-b02), Java HotSpot(TM) 64-Bit Server VM (build 25.333-b02, mixed mode)

(3) optional: Clampfit 10 (pClamp 10, Molecular Devices) for the convenient exploration of detection results (we also provide the Matlab app "CATHARSiS output viewer" as an alternative, see below)

Code provided

(1) CATHARSiS: a set of Matlab scripts and functions (bundled with (i) Fast Tiff Write v2.1 by R. Harkes and (ii) a modified version of UFARSA by V. Rahmati). Note: After unzipping, add all sub-folders to the Matlab path.

(2) returnImageJROIcoordinates.ijm: an ImageJ/Fiji macro that generates a tab-delimited TXT file containing the x-y coordinates of ImageJ ROIs (*.zip).

How to use

Step 1: Prepare files

Demo files

For testing CATHARSiS, demo files of types (1) – (4) (see below) can be conveniently generated by running the script generateDemoFiles.m. To this end, enter generateDemoFiles in the Matlab command line and press <Enter>. This will generate a new subfolder Demo_<timestamp> containing all files required for running CATHARSiS.

Here, we simulate spike-induced CaTs in two ring-shaped cells that partially overlap in space. Fluorescence signals of each cell are contaminated by (I) signals originating from the partially overlapping second cell, (II) spatially homogenous fluorescence changes mimicking axon-based neuropil activity and (III) a low level of Poissonian noise. Overlap can be varied using the parameter distrange in generateDemoFiles.m.

(1) a TIF file containing the Ca^{2+} imaging data

Important

CATHARSiS assumes that the stack has been properly registered (aligned).

(2) a tab-delimited TXT file specifying the coordinates (x-y pairs) of all individual pixels of all ROIs

- An ImageJ/Fiji macro (returnImageJROIcoordinates.ijm) is provided to convert ImageJ ROIs (*.roi or *.zip) to a TXT file:
  - How to use:
    Start Fiji > Plugins > Macros > Edit ... > Open returnImageJROIcoordinates.ijm > Ctrl+R > Select ImageJ ROI *.zip file in the file selection dialog box.
  - The TXT file containing ROI coordinates will then be saved in the same directory.

- The file format is as follows:

row_ROI 1	col_ROI 1	row_ROI 2	col_ROI 2	row_ROI 3	col_ROI 3
48	329	49	351	57	371
48	330	49	352	57	372
...	...	...	...	...	...
69	343	NaN	NaN	NaN	NaN

Important

Please use the Fiji and Java versions specified above, if you encounter any problem with this macro.

(3) a tab-delimited TXT file specifying intervals ('drift periods') that shall be excluded from analysis (e.g. reflecting drift due to animal movements)

- The file format is as follows (onset frames in the first column, offset frames in the second column):

100	200
500	800
1700	1900

Important
If no drift interval shall be excluded from analysis, an empty TXT file must be provided.

(4) a tab-delimited TXT file specifying file onset frames

- If the TIF file was obtained by merging several (discontinuous) recordings, CATHARSiS needs to be informed about the frame numbers corresponding to the first frame in each of them.
- The file format is as follows:

1
1501
3001

Important
If the TIF file represents a continuous recording, the TXT file only contains a '1' (that is, only the 1st frame represents a recording/file onset).

(5) optional: a MAT file specifying the final ΔF_{xy} templates, which were generated by a previous execution of CATHARSiS on the same ROIs

- This can be useful if you intend to compare ROIs in a BEFORE-AFTER setting. In this case, you might wish to use the same set of active cell templates for both conditions. To this end, (1) run CATHARSiS on the BEFORE dataset (TIF file) and (2) use the same templates for the AFTER dataset (TIF file) by specifying the MAT output file generated in the first step. See the parameter options.TemplatesExist below.
- This can also be used to re-apply UFARSA (with different detection settings) to D(t) without repeating the manual step of selecting candidate ΔF_{xy} templates.

Step 2: Specify parameters

Important
Parameters are specified in the Section "Specify parameters" in the script runCATHARSiS.m.

Parameters I: Template Matching

Parameter	Default *	Comment
options.firstframe	0	Set to 1, if the first frame in the TIF stack is NOT the first frame acquired, but e.g. an average used for stack alignment. Otherwise set to 0.

options.smooth	0	Set to 1 to apply Savitzky-Golay smoothing to $F_{xy}(t)$ and mean F(t). Otherwise set to 0.
options.Ntemplates	8	Number of candidate templates. A GUI is provided within CATHARSiS that allows the user to select those candidates that shall be used for computing the final template. Accepted range is 1 to 15.
options.Ntemplates_min	1	Minimum number of user-selected candidate templates for inclusion of that ROI. Range: 1 ≤ options.Ntemplates_min ≤ options.Ntemplates.
options.frameextension	5 **	Number of frames following each mean ΔF peak to be used for the generation of candidate templates (by averaging).
options.mindist	5 **	Minimum distance [frames] between the time points of any two ΔF peaks considered for candidate templates.
options.SG_length	6 **	Length [frames] of Savitzky-Golay smoothing window (applied to $F_{xy}(t)$ and mean $\dot{F}(t)$).
options.SG_order	2	Order [] of Savitzky-Golay smoothing (applied to $F_{xy}(t)$ and mean $\dot{F}(t)$).
options.smooth_CandCrit	1	Set to 1 to smooth the mean $\dot{F}(t)$ for the generation of candidate templates. Otherwise set to 0.
options.window_F0	500 **	Length [frames] of Savitzky-Golay smoothing window for computing the baseline $F0_{xy}(t)$. [] – use all frames.
options.F0_correct	–0.4	If the minimum ΔF/F0(t) for a given ROI is lower than options.F0_correct, F0(t) is instead calculated as the percentile specified by options.F0_percentile. Explanation: For highly active cells, oversmoothing (see above) might systematically overestimate F0(t).
options.F0_percentile	20	If the minimum ΔF/F0(t) is lower than options.F0_correct, F0(t) is instead calculated as the percentile specified by options.F0_percentile.
options.ROIexpansion	(0)	Euclidean distance [pixels] by which ROIs will be expanded. 0 – no expansion. Note: It might be useful that the final ROIs are somewhat larger than the actual cells, as the transition from (the brighter) cytosol to (the dimmer) neuropil has a high contrast. If imported ROIs correspond to the exact outlines of the cells, ROI expansion may be applied here without generating novel ROIs before.

options.TemplatesExist	0	Set to 1, if existing templates from a previous execution of CATHARSiS on the same ROIs shall be used. Otherwise set to 0.
options.SaveOptVars	0	Set to 1, if a number of optional variables are to be saved. Otherwise set to 0.

* Best values depend on many parameters, including sampling frequency (here, 10 Hz), type of indicator (here, GCaMP6s), spatial resolution etc.

** The given 'default' values for these parameter relate to the kinetics of somatic GCaMP6s signals and a frame rate of 10 Hz. Rule of thumb: For a frame rate *f*, multiply these values with a factor of $\frac{f}{10\,Hz}$ and round them up to the next integers.

Parameters II: UFARSA

Parameter	Default *	Comment
opt.scale_NoiseSTD	(2.5) *	Leading-threshold scaling constant. This is the main detection parameter in UFARSA. Rule of thumb: Decreasing the value of opt.scale_NoiseSTD increases sensitivity (recall) at the expense of lower specificity (precision). (Note: A lower bound for the leading threshold can be set via the parameter opt.min_leading_amp in the internal_parameters.m file of UFARSA.)
opt.remove_drifts	1	1: remove slowly varying drifts, 0: skip
opt.remove_posDeflections	0	1: apply large-impulse (deflection) removal step, 0:skip
opt.remove_negDeflections	1	1: remove large short-lasting negative deflections, 0: skip
opt.demerging	1	1: apply the demerging step, 0: skip

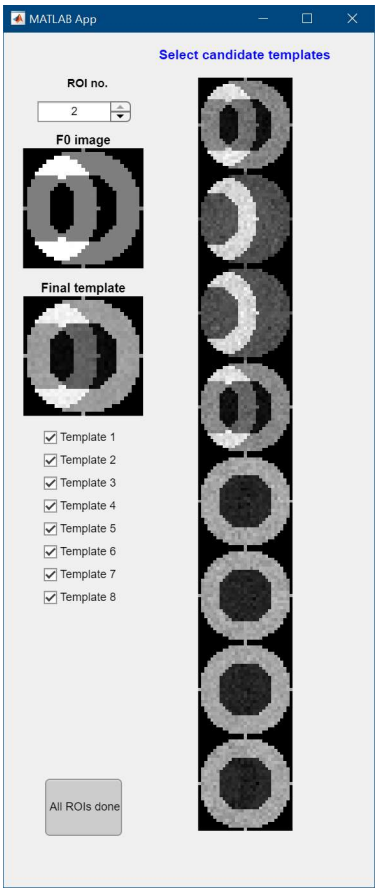
* Best values depend on many parameters, including sampling frequency (here, about 10 Hz), type of indicator (here, GCaMP6s), spatial resolution etc.

Important
For additional parameters related to UFARSA, see the internal_parameters.m file. There, in particular, a lower bound for the leading threshold can be set via the parameter opt.min_leading_amp. Depending on data characteristics, this might reduce false positive detection results.

Step 3: Run CATHARSiS

(1) Enter runCATHARSiS in the Matlab command line and press <Enter>.
(2) Specify the input files listed under step 1 (see above) in the file selection dialog boxes.
(3) A GUI opens that allows the user to select those candidates that shall be used for computing the final template. A preview of the final template is provided based on the currently selected candidates. Once you have finished the selection for all ROIs, press the 'All ROIs done' button (Fig. 1).

A



B

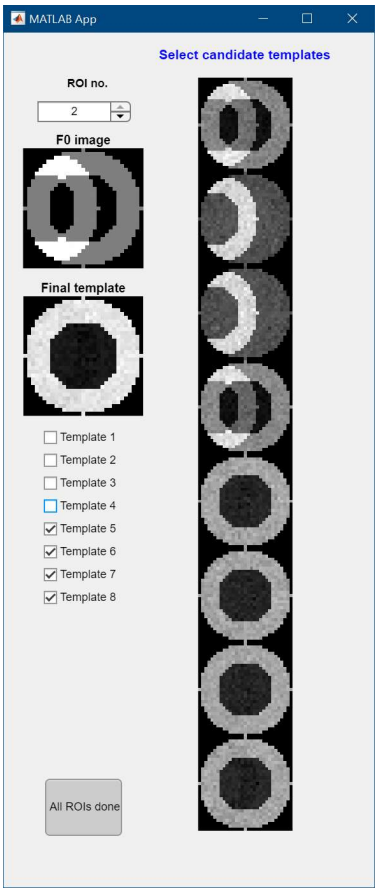


Fig. 1. GUI for the selection of candidate templates in CATHARSiS.

(A) All candidates are selected. Note that candidates #1–#4 are contaminated by an optically overlapping cell. In candidates #2 and #3, the current cell of interest is in fact inactive. Including these candidates would result in a sub-optimal final template (lower-left image).
(B) After de-selecting candidates #1–#4, the final template is an adequate representation of the active cell.
(A–B) Use the spinner on the top-left to switch between ROIs. After completing the (de-)selection of candidate templates for all ROIs, press the 'All ROIs done' button to finalize event detection. The number of candidates shown in the GUI (here, 8) is set by options.Ntemplates (see Step 2).

Important

If options.TemplatesExist is set to 1, the user is requested to select the MAT file that specifies the existing ΔF_{xy} templates generated by a previous execution of CATHARSiS on the same ROIs. In this case, the GUI-based manual selection of ΔF_{xy} templates (Fig. 1) is omitted.

Step 4: Evaluate the outputs

CATHARSiS generates two output files:

(1) a tab-delimited TXT file for the convenient exploration of detection results using Clampfit 10 (pClamp 10, Molecular Devices):

The file can be imported into Clampfit via drag & drop (the number of signals must be set to 5). Each ROI will be shown as a sweep with the following 5 signals:

Signal	Comment	Corresponding variable in *.mat (k denotes the ROI number)
Detected CaT onsets	1 – onset, 0 – no onset	UFARSA_output{k}.output_UFARSA.eTrain_dem
Detection criterion D(t) (processed by UFARSA)		UFARSA_output{k}.output_UFARSA.fluors.afterSmoothing
Detection criterion D(t) (raw)		UFARSA_output{k}.output_UFARSA.fluors.original
Mean ΔF/F0(t)		DF_F0_perROI(:, k)
Drift periods	1 – drift, 0 – no drift	driftindex

Note
NaNs in the above Matlab variables are replaced by zeros, as Clampfit cannot handle NaNs. For all frames belonging to a drift period, the values of signals 1–4 are set to zero.

Note
Alternatively, the above signals can be inspected using the output viewer app (Fig. 2). The app requires that the variable Output (generated by CATHARSiS and saved in the MAT file, see below) is present in the base workspace. To run it, enter OutputViewer in the command line and press <Enter>.

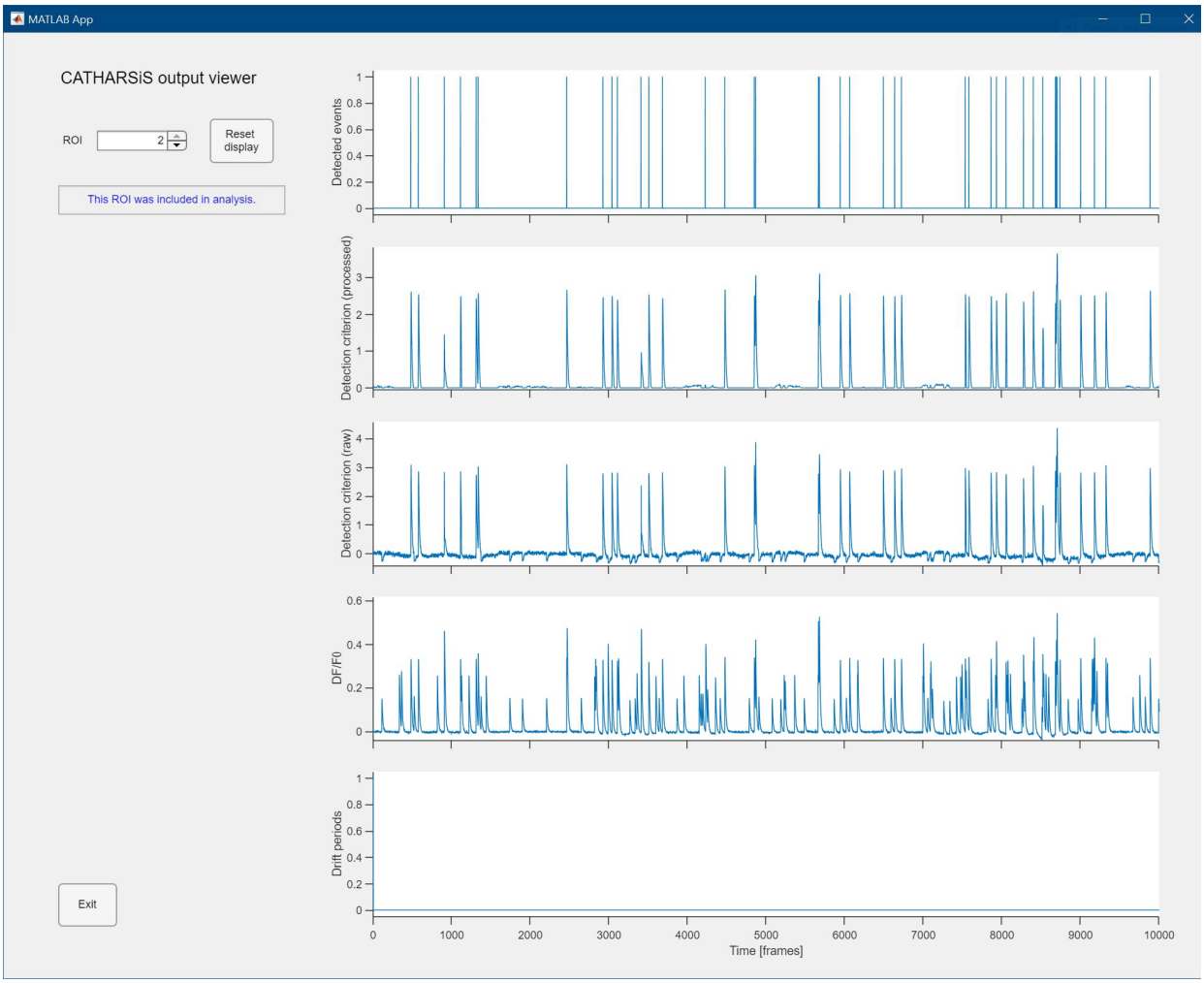


Fig. 2. CATHARSiS output viewer. Use the spinner on the top-left to switch between ROIs. Five signals are shown (from top to bottom): (1) detected CaT onsets (1 – onset, 0 – no onset); (2) D(t) (processed by UFARSA); (3) D(t) (raw); (4) $\Delta F/F_0(t)$; (5) drift periods (1 – drift, 0 – no drift). Click on a plot for panning or zooming.

(2) a MAT file containing all detection results and parameter settings (most important outputs are highlighted in blue)

Variable name	Comment
options	Parameter values related to CATHARSiS (set in runCATHARSiS.m)
opt	Parameter values related to UFARSA (only those set in runCATHARSiS.m)
Results_perROI	Main summary of the detection results, separately for each ROI
▪ row 1: ROI number	ROI number
▪ row 2: F(xyt)	(empty by default)
▪ row 3: F0(xyt)	(empty by default)
▪ row 4: DF(xyt)	(empty by default)
▪ row 5: Rise peaks	Time points [frames] of the candidate templates
▪ row 6: Candidate templates (xyn)	Candidate templates
▪ row 7: Unused	(empty by default)
▪ row 8: User-selected templates (0 - deselected)	Binary vector indicating whether a candidate ΔF_{xy} template was selected (1) or deselected (0) by the user
▪ row 9: Final template (xy)	Final ΔF_{xy} template used for computing D(t)
▪ row 10: DF/F0(t) per ROI w/o expansion	Mean $\Delta F/F0$ (ROI expansion is not applied here)
▪ row 11: Detection criterion	Raw D(t)
▪ row 12: UFARSA: Demerged event train	Activity vector (1 – CaT onsets, 0 – no CaT onset, NaN – frame in drift period)
▪ row 13: UFARSA: leading_thr	Leading threshold computed and used by UFARSA
▪ row 14: UFARSA: smoothing_param	Smoothing parameter computed and used by UFARSA
driftindex	Binary vector of length Nframes (1 – frame within drift period, 0 – frame outside drift periods)
FileOnsetFrames	Frames corresponding to file onsets (as defined in the corresponding input TXT file)
finalFrames	Vector of final 'non-drift' frames (here, 'drift' includes: (i) drift periods defined by the corresponding input TXT file, (ii) frames corresponding to file onsets, (iii) each first frame after each drift period) *
finalROIs	Vector of ROIs for which a final ΔF_{xy} template was computed, i.e. ROIs included in event detection
Nframes	Total number of frames ('non-drift' and 'drift' frames)

non_finalROIs	Vector of ROIs for which a final ΔF_{xy} template was NOT computed, i.e. ROIs excluded from event detection
Nrois	Total number of ROIs
Output	Contains all data exported to a tab-delimited TXT file which can be used for the convenient exploration of detection results using Clampfit 10. Alternatively, you may use the CATHARSiS OutputViewer app (Fig. 2).
UFARSA_output	Generated by UFARSA. Partially redundant with some of the variables above. Important: UFARSA_output{1, k}.opt_out contains all parameter settings of UFARSA for the k-th ROI (which are set in or computed by runCATHARSiS.m or internal_parameters.m).

* We consider frames according to (ii) and (iii) as 'drift' since UFARSA tends to detect a CaT onset in these frames if there is a rise/decay of a CaT whose actual onset is *before* that frame. Considering these frames could otherwise lead to spurious synchronicity.

If options.SaveOptVars is set to 1 (see above), the following additional variables are additionally saved in the MAT file.

Variable name	Comment
coord	Coordinates (x-y pairs) of all individual pixels of all ROIs specified by the user in a tab-delimited TXT file
coord_extended	Coordinates (x-y pairs) of all individual pixels of all ROIs specified after (optional) ROI expansion (see options.ROIexpansion)
DF_F0_perROI_min	Minimum of mean $\Delta F/F0(t)$ for each ROI (see options.F0_correct)
F_perROI	F(t) for each ROI
ROIs_Prctile_F0	Binary vector indicating if F0(t) was calculated as the percentile specified by options.F0_percentile