The Journal of Open Source Software

SNAzzy: an image processing pipeline for investigating global Synchronous Network Activity

Carlos D. Paiva [1,2], **Alana J. Evora** [1,2], **Shirui Zheng** [1,2], **and Arnaldo Carreira-Rosario** [1,2¶]

1 Department of Biology, Duke University, Durham, NC 27708 **2** Department of Neurobiology, Duke University, Durham, NC 27708 ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor:

Submitted: 06 November 2025
Published: unpublished

Summary

Genetically encoded fluorescent indicators are powerful tools for monitoring biological processes in live samples (Lin & Schnitzer, 2016; Nakai et al., 2001). When combined with a large field of view, a single time-lapse recording has the potential to capture many specimens, facilitating high-throughput data collection. However, this approach generates large, multidimensional datasets that are challenging to process and analyze. We present `SNAzzy`, a Python package for studying synchronous network activity (SNA) in Drosophila embryos via high-throughput microscopy. SNA is a hallmark of developing nervous systems (Akin & Zipursky, 2020; Blankenship & Feller, 2009; Wu et al., 2024), often studied using genetically encoded calcium indicators to monitor neural activity in vivo. `SNAzzy` processes and analyzes time-lapse datasets taken from live samples using fluorescent widefield microscopy. Each dataset contains dozens of individual specimens in the same field of view and thousands of time points. The software offers individual specimen cropping for optimization of storage and processing, adaptive regions of interest for quantification of fluorescence and changes in morphology over time, a custom peak detection algorithm, and a graphical user interface for data visualization, curation, and dataset comparison. This tool can be readily applied to analyze fluorescent intensities in time-lapse microscopy experiments that involve simultaneous imaging of multiple samples, particularly small-sized specimens (Avasthi et al., 2023; Donoughe et al., 2018).

Statement of need

During synchronous network activity (SNA), many neurons fire simultaneously, generating waves of activity that span across large portions of the nervous system (Akin & Zipursky, 2020; Blankenship & Feller, 2009; Wu et al., 2024). In Drosophila embryos, SNA typically lasts 4 hours, during which the nervous system undergoes a stereotyped morphological change via ventral nerve cord condensation (Carreira-Rosario et al., 2021; Crisp et al., 2008; Karkali et al., 2022). To gain an understanding of SNA, it is essential to quantify waves of activity in the nervous system while also tracking morphology as a proxy for neurodevelopment. For these reasons, we combine a commonly used genetically encoded calcium indicator (GECI) that reports neural activity with a structural fluorophore (Carreira-Rosario et al., 2021). The structural fluorophore signal remains stable, independent of neural activity, making it suitable for continuous tracking morphology of the nerve cord. To record many embryos during SNA, we use a wide-field fluorescence microscopy system that captures the GECI and structural fluorophore signal of dozens of developing embryos for over 5 hours.

We were unable to find a tool designed for widefield microscopy that rapidly processes multiple specimens, quantifies levels of fluorophore activity, and incorporates a peak-finding algorithm suitable for global calcium traces. `SNAzzy` is designed to investigate global levels of neural

activity across multiple developing embryos simultaneously.

Tracking of multiple "adaptive ROIs"

To the best of our knowledge, there are no other packages that provide functionality for automated parsing of raw images of many live specimens into activity and morphological quantifications. Other studies have employed manual selection of regions of interest (ROIs) and used static ROIs (Akin & Zipursky, 2020; Ardiel et al., 2022; Carreira-Rosario et al., 2021; Menzies et al., 2024). Manual selection often generates imprecise ROIs, which can lead to inaccurate quantifications, and is also cumbersome and prone to human error. Static ROIs are not reliable for detecting the fluorescent signal of live specimens that change in morphology and move while imaging. `SNAzzy` fills these gaps as an accessible pipeline for the automated analysis of multiple live samples in parallel. The pipeline generates an "adaptive ROI" that changes frame-by-frame for each specimen. This enables the accurate tracking of fluorescence intensity as well as changes in tissue morphology or size. `SNAzzy`'s design provides an automated, modular, and fully auditable workflow, and ultimately contributes to more reproducible and comparable results across experiments.

Capturing global Calcium dynamics

To the best of our knowledge, there are no open-source packages that provide tools for performing automated data analysis and quantification of global calcium dynamics. Most open-source tools available for analyzing neural activity using GECI focus on segmenting individual neurons within a single specimen. `CaImAn` (Giovannucci et al., 2019), and `Suite2p` (Pachitariu et al., 2016) are among the most widely used. These packages detect calcium dynamics and use individual neuron statistics to perform spike inference, but do not offer direct peak detection on the calcium signal. Furthermore, they are optimized for two-photon microscopy as opposed to wide-field microscopy. `SNAzzy` provides a series of automated analyses and quantifications to analyze global calcium levels in time-series acquired with widefield microscopes.

The Journal of Open Source Software

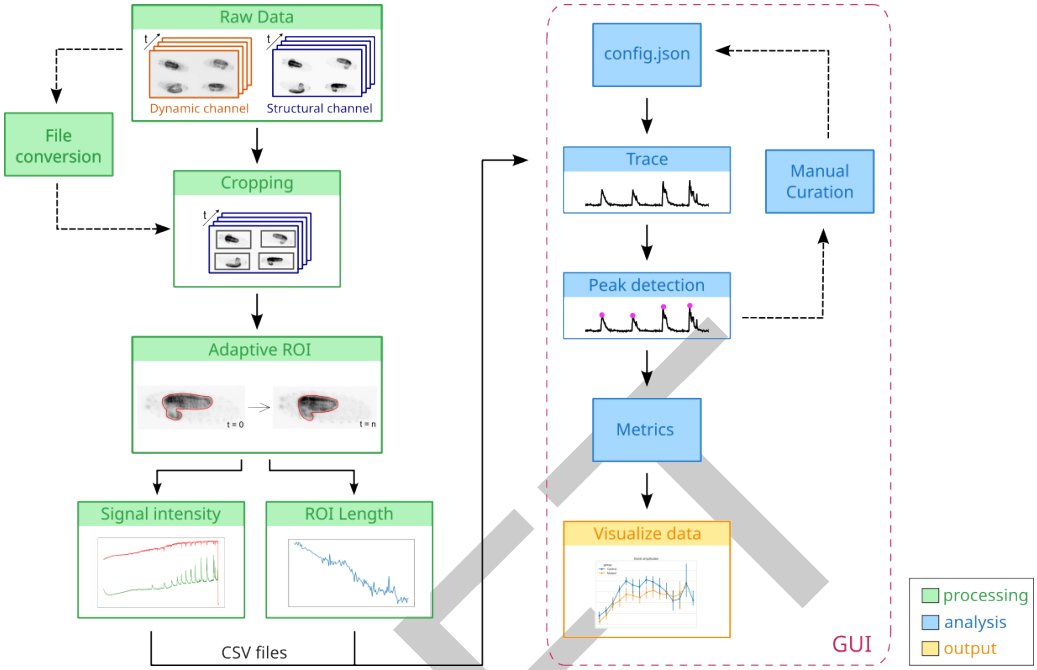

Figure 1: Schematic of the SNAzzy pipeline. Time-lapse taken from fluorescent widefield microscopes (raw data) enters the processing stage (green). The processing stage outputs two types of CSV files: time series of signal intensities from each recorded channel and ROI length. CSV files enter the analysis stage (blue) to generate normalized fluorescent traces and detect peaks along with other signal processing metrics. These initial traces can be visualized to curate the data. Curation generates a configuration file that works as metadata across platforms and users. Curated data can be reanalyzed and used to visualize final data and compare across groups (yellow). Analysis and output stages are performed in the GUI (red dashed box), along with other metrics. Dashed arrows indicate optional steps.

Pipeline Description

The initial input for SNAzzy Figure 1 is raw time-lapse imaging data containing multiple embryos. Each embryo expresses a GECI (dynamic fluorophore) and a structural fluorophore. Fluorophores are imaged in different optical channels.

The first pipeline step converts the raw data to TIF format, thereby avoiding compatibility issues that may arise when parsing different proprietary formats Figure 1. All embryos are then segmented using histogram equalization, followed by intensity threshold binarization (Otsu, 1979). Boxes surrounding the segments are cropped into individual time-lapses for each embryo. Cropping results in a substantial memory reduction, as most background pixels are removed, with cropped images typically accounting for around 40% of the original size.

The next step is to process each individual specimen. First, the ROI, which in our case is the entire central nervous system (CNS), is defined by binarizing the structural channel and selecting the largest connected component. This process is repeated at every time point to generate an "adaptive ROI". From these adaptive ROI, the average signal intensity for both channels is extracted. The results are saved as CSV files and are the basis for downstream analysis.

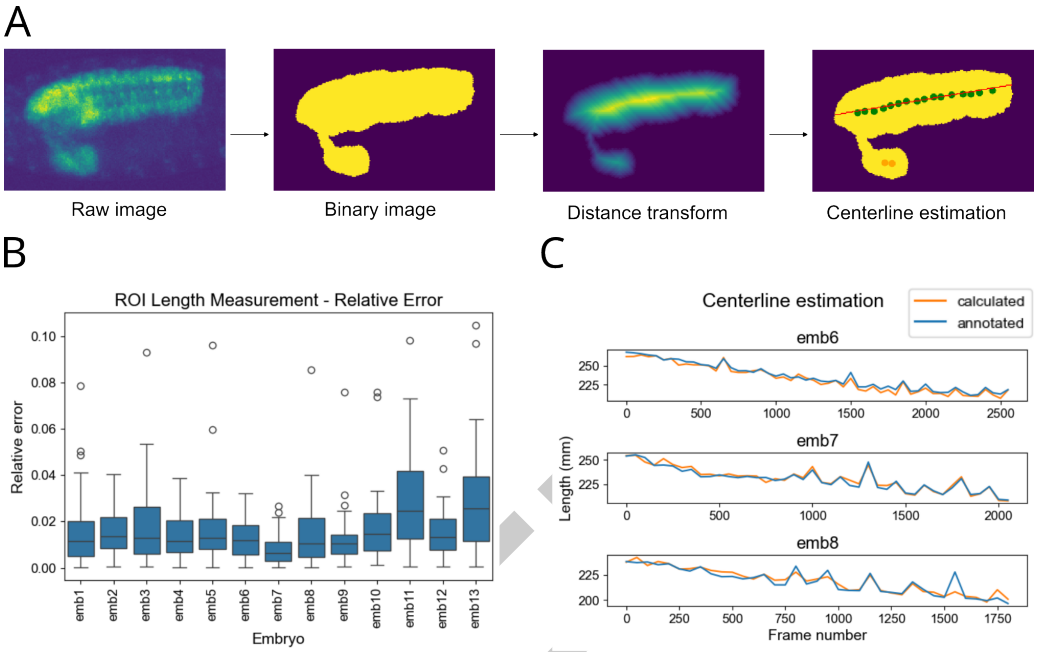

Figure 2: ROI length measurement algorithm and validation. A) Steps to calculate the ROI length. The ROI length is calculated by estimating the centerline (red line) using points of maximum (dots) in the distance transform, followed by RANSAC to ignore outliers (orange dots). B) Validation of the method as relative error (measured - annotated) / annotated. Each whisker bar summarizes the relative error for frames taken at intervals of 50 timepoints. C) Comparison of absolute values over a time series for three representative embryos.

The ROI is also used to measure the length of the CNS Figure 2. Drosophila embryo CNS length serves as an internal proxy for neurodevelopmental stages, enabling more accurate comparisons across embryos (Carreira-Rosario et al., 2021). The CNS length is calculated by centerline estimation. First, a distance transform is applied to the binarized image, and local maxima points are detected. Depending on the embryo's orientation, some points may be part of the brain lobes and must be filtered out to accurately measure the CNS length. To obtain a robust centerline estimate that can ignore outliers, we use RANSAC (Fischler & Bolles, 1981) over the local maxima points and measure the overlap between the fitted line and the binary image. CNS length is also detected frame by frame and exported as a CSV file Figure 1.

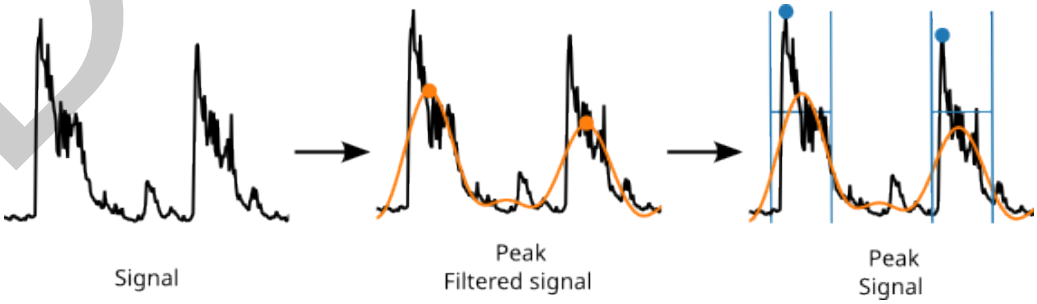

Figure 3: Peak detection algorithm. A low-pass filter (orange line) is applied to the ΔF/F signal (black line) to remove fast transients. The peak in the filtered signal (orange dot) is then ported back to the ΔF/F (blue dot) signal by selecting the leftmost peak within a search window (blue lines).

The package utilizes average signal intensity measurements to calculate ΔF/F traces and peaks. For ΔF/F, we first calculate the ratiometric signal (dynamic signal / structural signal) and

then its baseline, which is defined as the average of the N lowest values within a sliding window. The generated $\Delta F/F$ traces contain long-duration bouts of activity with superimposed fast transients Figure 3. The former represents the bursts of activity and is the most relevant for the initial analysis. To mark only these more prolonged bouts, we apply a low-pass frequency filter to omit transients. Peaks in the filtered trace are detected using SciPy (Virtanen et al., 2020). Finally, the detected peaks are ported to the original $\Delta F/F$ signal.

Results can be visualized and curated in a graphical user interface (GUI) implemented in `PyQt6` Figure 4. During curation, researchers can modify data analysis parameters, which are persisted in a JSON configuration file and utilized by the core analysis code across different machines and users. Finally, a large number of different metrics and representations derived from $\Delta F/F$, CNS length, and peaks can be visualized and plotted using the GUI. These include SNA onset, burst duration and spectrograms, among others.

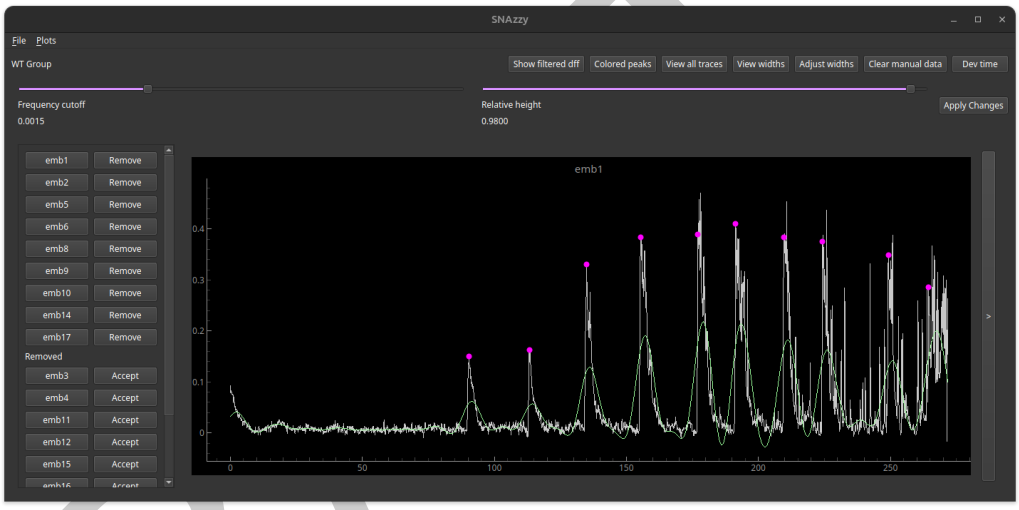

Figure 4: GUI for data validation, curation, visualization and plotting. Initial GUI screen. A $\Delta F/F$ trace (white) and the corresponding peaks (magenta dots) are shown. The low-passed signal (green line) is used as a reference to determine peaks. The GUI enables the modification of analysis parameters, visualization of data, and comparison of metrics across groups of experiments, as well as manual adjustment of peak data.

Conclusion

In conclusion, genetically encoded fluorescent indicators and microscopy systems are evolving rapidly, increasing the data acquisition throughput. Custom open-source tools are needed to handle such large data files. `SNAzzy` addresses this by offering an automated, scalable, and user-friendly platform for analyzing synchronous network activity in developing embryos. As an open and versatile solution, `SNAzzy` offers tools for a broader range of applications in time-lapse fluorescence imaging across diverse biological systems.

Acknowledgments

We acknowledge Newton PenkoffLidbeck and D. Berfin Azizoglu for feedback on the manuscript. This work was partially funded by NINDS and the BRAIN initiative (R00NS119295).

References

Akin, O., & Zipursky, S. L. (2020). Activity regulates brain development in the fly. *Current Opinion in Genetics & Development*, *65*, 8–13. https://doi.org/10.1016/j.gde.2020.04.005

Ardiel, E. L., Lauziere, A., Xu, S., Harvey, B. J., Christensen, R. P., Nurrish, S., Kaplan, J. M., & Shroff, H. (2022). Stereotyped behavioral maturation and rhythmic quiescence in c. Elegans embryos. *eLife*, *11*. https://doi.org/10.7554/eLife.76836

Avasthi, P., Essock-Burns, T., Garcia, I., Galo, Gehring, J., Matus, D. Q., Mets, D. G., & York, R. (2023). *Gotta catch 'em all: Agar microchambers for high-throughput single-cell live imaging*. https://doi.org/10.57844/arcadia-v1bg-6b60

Blankenship, A. G., & Feller, M. B. (2009). Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nature Reviews. Neuroscience*, *11*(1), 18–29. https://doi.org/10.1038/nrn2759

Carreira-Rosario, A., York, R. A., Choi, M., Doe, C. Q., & Clandinin, T. R. (2021). Mechanosensory input during circuit formation shapes drosophila motor behavior through patterned spontaneous network activity. *Current Biology: CB*, *31*(23), 5341–5349.e4. https://doi.org/10.1016/j.cub.2021.08.022

Crisp, S., Evers, J. F., Fiala, A., & Bate, M. (2008). *The development of motor coordination in drosophila embryos*. *3717*, 3707–3717. https://doi.org/10.1242/dev.026773

Donoughe, S., Kim, C., & Extavour, C. G. (2018). High-throughput live-imaging of embryos in microwell arrays using a modular specimen mounting system. *Biology Open*, *7*(7), bio031260. https://doi.org/10.1242/bio.031260

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381–395. https://doi.org/10.1145/358669.358692

Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., Taxidis, J., Najafi, F., Gauthier, J. L., Zhou, P., Khakh, B. S., Tank, D. W., Chklovskii, D. B., & Pnevmatikakis, E. A. (2019). CaImAn an open source tool for scalable calcium imaging data analysis. *eLife*, *8*. https://doi.org/10.7554/eLife.38173

Karkali, K., Tiwari, P., Singh, A., Tlili, S., Jorba, I., Navajas, D., Muñoz, J. J., Saunders, T. E., & Martin-Blanco, E. (2022). Condensation of the drosophila nerve cord is oscillatory and depends on coordinated mechanical interactions. *Developmental Cell*, *57*(7), 867–882.e5. https://doi.org/10.1016/j.devcel.2022.03.007

Lin, M. Z., & Schnitzer, M. J. (2016). Genetically encoded indicators of neuronal activity. *Nature Neuroscience*, *19*(9), 1142–1153. https://doi.org/10.1038/nn.4359

Menzies, J. A. C., Chagas, A. M., Baden, T., & Alonso, C. R. (2024). A microRNA that controls the emergence of embryonic movement. In *eLife*. https://doi.org/10.7554/elife.95209.2

Nakai, J., Ohkura, M., & Imoto, K. (2001). A high signal-to-noise ca(2+) probe composed of a single green fluorescent protein. *Nature Biotechnology*, *19*(2), 137–141. https://doi.org/10.1038/84397

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L. F., Dalgleish, H., Carandini, M., & Harris, K. D. (2016). Suite2p: Beyond 10,000 neurons with standard two-photon microscopy. In *bioRxiv* (p. 061507). bioRxiv. https://doi.org/10.1101/061507

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D.,

Burovski, E., Peterson, P., Weckesser, W., Bright, J., Walt, S. J. van der, Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Contributors, S. 1.0. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wu, M. W., Kourdougli, N., & Portera-Cailliau, C. (2024). Network state transitions during cortical development. *Nature Reviews. Neuroscience*. https://doi.org/10.1038/s41583-024-00824-y