Srujan Yamali

302-509-8614 | srujanyamali@berkeley.edu | linkedin.com/in/srujanyamali | github.com/Srujyama | Srujyama.vercel.app

Education

University of California, Berkeley May 2027
B.S. Computer Science *GPA: 4.0/4.0*

Experience

Data Science Intern November 2023 – August 2025
University of Pennsylvania | Children's Hospital of Philadelphia - Moustafa Lab *Philadelphia, PA*

- Designed Carpet Cleaned Changepoints, a high-performance Python tool using `ruptures`, `KDTree`, and multiprocessing to identify recombination events in **75,000+** microbial genomes (**37 TB**).
- Optimized a **time series analysis** pipeline for bacterial proteomic datasets using `kernel-based changepoint detection` (KernelCPD) to identify statistically significant shifts in protein signal distributions.
- Engineered a parallelized framework leveraging Python's multiprocessing to accelerate detection across genomic regions, followed by region-wise t-test comparisons and Jaccard-similarity logic for adjacent-region merging.
- Generated Matplotlib heatmaps and line plots, exporting batch reports for large-scale comparative genomics.

Machine Learning Research Intern March 2024 – August 2024
Cornell University - Hein Lab *Remote*

- Applied deep learning techniques, including `YOLO-based object detection`, to automate the identification and tracking of individual and group fish behaviors due to predation from field video data achieving **85%+** accuracy in multi-class fish behavior prediction across **500+** hours of field video footage.
- Assisted in inferring decision-making rules using `stochastic models` and dynamic behavioral sequences.

Computer Vision Research Intern June 2023 – August 2024
University of Delaware – Shao Lab *Newark, DE*

- Developed a PyQt6/OpenCV application to automate analysis of **730 GB** of Drosophila video, reducing manual annotation time by **90%** saving hundreds of hours
- Designed an ROI-tracking engine using contour detection, blob tracking, and centroid calculations to monitor thigmotactic behavior and wall-following dynamics during mating trials achieving **99.7%** accuracy.
- Engineered spatial analysis logic to quantify centrophobism via time-in-center metrics, emitting real-time signals to a GUI overlay for behavioral classification and gender-role tracking.
- Implemented data smoothing, latency modeling, and correlation testing (Pearson, Mann–Whitney U) to statistically link mating duration with anxiety-like spatial behaviors.

Projects

Real-Time Behavioral Detection and Tracking | *Python, OpenCV, PyQt6, Pandas* `DrosophilaCV`

- Built an end-to-end engine combining real-time object tracking for mating/thigmotaxis behaviors (OpenCV), a PyQt6 GUI for video annotation and CSV export, plus optimized NumPy/Pandas pipelines for batch trajectory processing, spatial metrics computation, and behavior event labeling.

Genetic Changepoint Heat-Map Engine | *Python, ruptures, Scikit-learn, Matplotlib* `RedCarpet`

- Created a high-performance engine for detecting statistical changepoints in large genomic sequences using SciPy, performing matrix operations and clustering via NumPy, and visualizing results through Matplotlib-based heatmaps, scalable to high-dimensional datasets.

Publications & Conferences

A. Moustafa, E. Theiller, A. Lal, S. Yamali, A. Feder, A. Narechania, P. Planet. *Redcarpet: A Tool for Rapid Recombination Detection in Staphylococcus aureus.* 19th ISSSI, 2024. **icmsmeetings.eventsair.com/isssi-2024/speakers**

R. Oliver, S. Yamali, S. Knox, T. Dadyala, L. Shao. *High-Throughput Behavioral Assay Unveils Female Courtship in Drosophila.* Proceedings of the International Behavioral and Neural Genetics Society, 2024. **IBANGS.org**

R. Oliver, S. Yamali, S. Knox, T. Dadyala, L. Shao. *High-Throughput Behavioral Assay Unveils Female Courtship in Drosophila.* Sexually Dimorphic Circuits and Behaviors, Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, 2024.

Technical Skills

Languages: Python, C++, JavaScript, HTML/CSS, R, Java, Golang, SQL
Libraries: Pandas, OpenCV, Ruptures, SciPy, Scikit-Learn, Matplotlib, Seaborn, PyQt6, TensorFlow, Keras, PyTorch
Tools/Technologies: Git, Docker, Kubernetes, PostgreSQL, MySQL