

Datasets for Benchmarking Behavior Segmentation

Large-scale datasets that can be used to evaluate behavior segmentation models across different settings (video-level/frame-level, single/multi-agent, different organisms)

Dataset	Task	Setting	Organism
<u>CRIM13</u>	Frame-level Behavior Classification	Videos from 2 views, with 13 annotated social behaviors	Mouse
<u>Fly vs. Fly</u>	Frame-level Behavior Classification (also has 2D poses)	Videos & trajectory, with 10 annotated social behaviors	Fly
<u>CaIMS21</u>	Frame-level Behavior Classification (also has 2D poses)	Videos & trajectory, with 10 annotated social behaviors	Mouse