Using Deep Learning to Quantify Mouse Orofacial Behavior

Sydney Bergstrom[1], Kerry Nix[2], Timothy Machado[2]

[1]SAS 2026, [2]Perelman School of Medicine, Department of Neuroscience

1. How Can We Quantify Animal Behavior?

Dakota Digital Review

- Computer vision enables computers to understand visual information and teaches them to make decisions based on what they see
- Deep learning creates neural networks that learn features of visual data to identify body parts and label them[1]
- Social Leap Estimates Animal Poses (SLEAP)[2] is one of the most advanced deep learning frameworks for animal pose tracking that aims to quantify animal behavior

2. Using Deep Learning to Track Body Movement

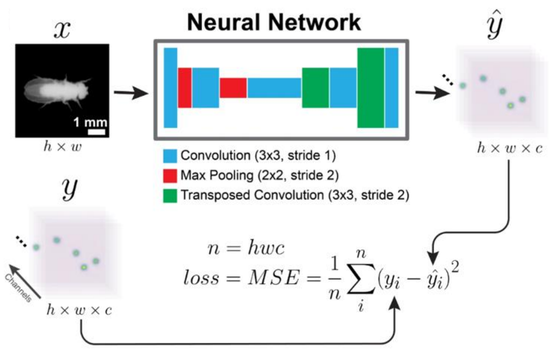
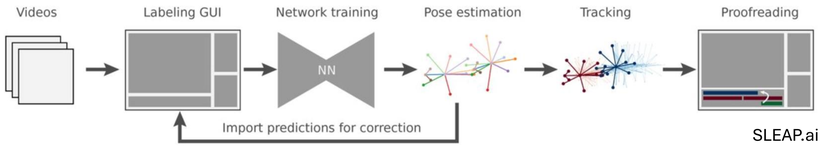


Diagram of the SLEAP neural network architecture.

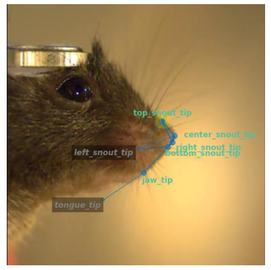


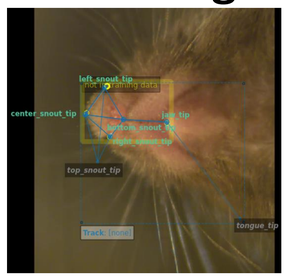
SLEAP.ai

Diagram of the steps for creating a project in SLEAP. You first import videos; then define the animal skeleton and select frames to label; then train a model using those labeled frames until desired model accuracy; then apply the model to predict poses in unlabeled frames; and finally proofread labeled frames and export for analysis

Thank you PURM, Penn Medicine, and everyone at the Machado lab for your support!

3. SLEAP Model Training and Data Processing



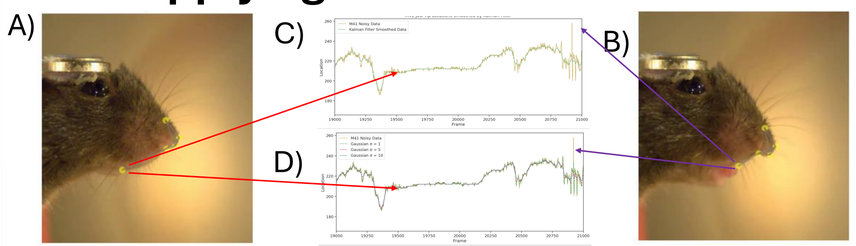


The same 7 nodes and 10 edges were labeled across frames in videos of 11 mice (175 frames in the side lab model and 285 frames in the lower lab model).

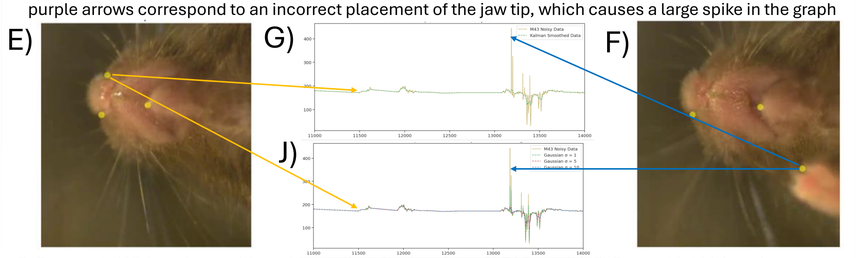
Side Video With Predicted Points

Lower Video With Predicted Points

4. Applying Models to New Mice

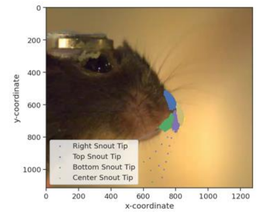


A) Frame 19,504 from side video with predicted points from SLEAP model. B) Frame 20,921 from side video. C) Jaw tip locations of mouse M41 put through a Kalman filter. D) Jaw tip locations of mouse M41 put through a Gaussian filter. The red arrows correspond to a correct placement of the jaw tip on graphs C) and D). The purple arrows correspond to an incorrect placement of the jaw tip, which causes a large spike in the graph



E) Frame 11,497 from lower video with predicted points from SLEAP model. F) Frame 13,186 from lower video. G) Left snout tip locations of mouse M43 put through a Kalman filter. H) Left snout tip locations of mouse M43 put through a Gaussian filter. The yellow arrows correspond to a correct placement of the left snout tip on graphs G) and J). The blue arrows correspond to an incorrect placement of the jaw tip, which causes a large spike in the graph.

5. Filtering Reduces the Frequency of Errors



Right snout tip (purple), top snout tip (blue), bottom snout tip (green), and center snout tip (yellow) locations across the whole video plotted on a still image of the mouse for the side lab model.

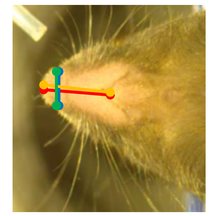
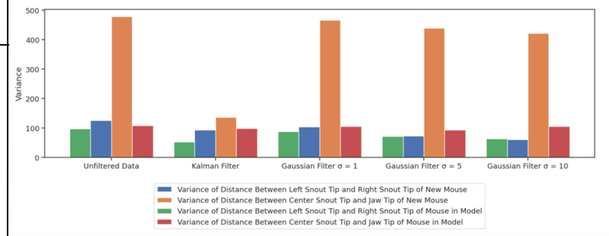
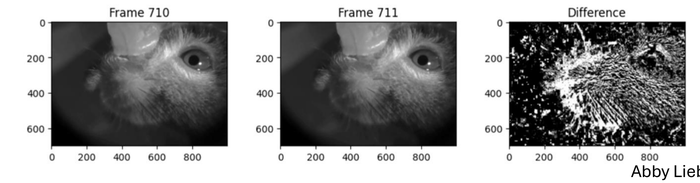


Figure of the distances measured. The orange and red lines are the distance between the center snout and jaw tip. The green and blue lines are the distance between the left and right snout tips.



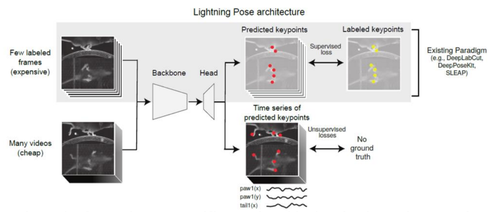
Variance of distance for two different sets of points for a held-out mouse trained on the lower lab model compared to the same points for a mouse included in the lower lab model training set and trained on it.

6. Conclusions and Future Directions



Abby Lieberman

Motion energy analysis[3] measures movement in regions of interest by measuring the absolute value in differences in pixel values in consecutive frames.



Lightning Pose[4] is another platform that quantifies animal behavior by using semi-supervised learning

- Large spikes in graphs probably correspond to incorrect node placement
- Kalman filtering smooths movements the most
- More frames should be labeled to reduce outliers
- Some nodes are tracked better than others
- Lightning Pose is another method to quantify behavior
- Anipose[5] does 3D tracking of animal poses

7. References

[1]Pereira, T.D., Shaevitz, J.W. & Murthy, M. Quantifying behavior to understand the brain. *Nat Neurosci* **23**, 1537–1549 (2020). https://doi.org/10.1038/s41593-020-00734-z

[2]Pereira, T.D., Tabris, N., Matsliah, A. *et al.* SLEAP: A deep learning system for multi-animal pose tracking. *Nat Methods* **19**, 486–495 (2022). https://doi.org/10.1038/s41592-022-01426-1

[3]Carsen Stringer *et al.* Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364** (2019). DOI: https://doi.org/10.1126/science.aav7893

[4]Biderman, D., Whiteway, M.R., Hurwitz, C. *et al.* Lightning Pose: improved animal pose estimation via semi-supervised learning, Bayesian ensembling and cloud-native open-source tools. *Nat Methods* **21**, 1316–1328 (2024). https://doi.org/10.1038/s41592-024-02319-1

[5]Karshchuk, P. *et al.* Anipose: A toolkit for robust markerless 3D pose estimation. *Cell Reports* **36** (2021). DOI: https://doi.org/10.1016/j.celrep.2021.109730