Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions (2020)

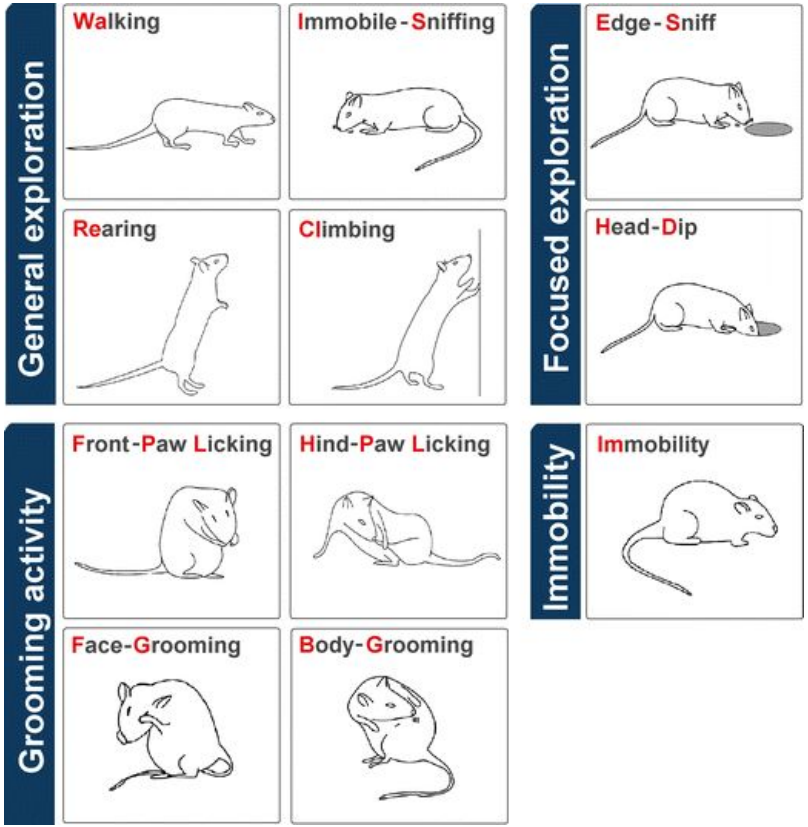
Sturman, O., von Ziegler, L., Schläppi, C. et al.
Neuroscience Center, ETH Zurich & University of Zurich, Switzerland

Sachin Salim

Draelos Lab: Paper Presentation
09 June 2023

Introduction

- **Objective:** Comparison of commercial platforms for behavior analysis such as EthoVision and TSE Multi Conditioning System with DeepLabCut (with post-hoc analysis)
- **Motivation:** Accurate behavioral analysis is crucial to study brain function in animals by gaining insights into the underlying neural mechanisms
- I'm interested in exploring the neural basis of behavior and constructing computational models using deep learning



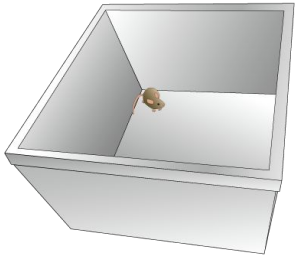
Note: All figures in this presentation are used from the original paper unless mentioned otherwise

researchgate.net

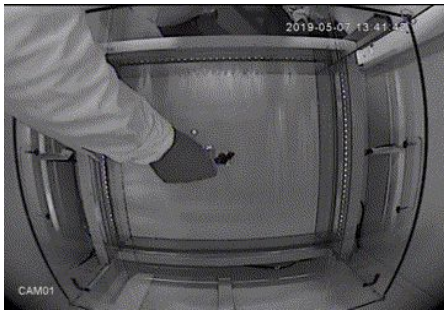
Methods

Behavioral tests

Open Field Test (OFT)

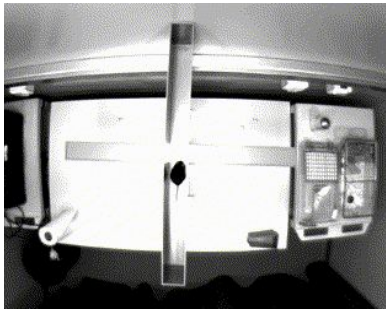


brainstuff.org



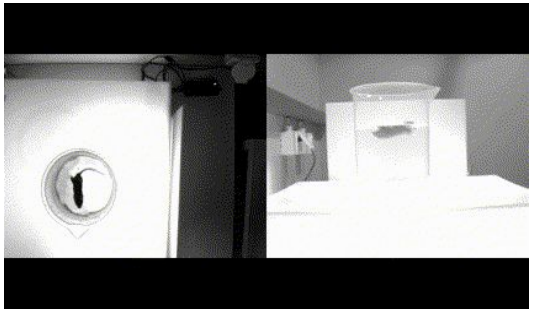
Elevated Plus Maze (EPM)

ugobasile.com



Forced Swim Test (FST)

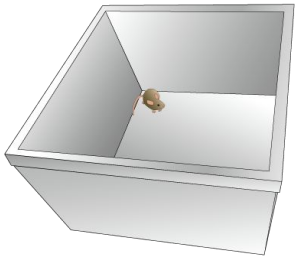
creative-biolabs.com



Video Source: zenodo.org

Behavioral tests

Open Field Test (OFT)



- Distance
- Time in center
- Supported rear
- Unsupported rear

Elevated Plus Maze (EPM)

- Distance
- Time in zones
  - Open/closed
  - Center
- Head dips

Forced Swim Test (FST)

- Distance
- Time floating

Ethological behaviors

What is it?

- behavior of animals under natural conditions, typically associated with emotional and disease states

Examples in mice

- rearing in OFT
- head dipping in EPM
- floating in FST

What do they signify?

- reduced exploration (rearing/head dipping) indicates anxiety
- floating has been linked to adaptive stress-coping behaviors

yourcounselling.ca

Capturing and Analyzing Videos

- Commercial Solutions
  - **Noldus EthoVision XT14**
  - **TSE Multi Conditioning System**

- Their Disadvantages
  - Expensive, more features require additional purchase
  - Lack flexibility to define and score specific behaviors of interest
  - Cannot be adapted to fit changing experimental needs
  - Suboptimal tracking ability
  - Poor sensitivity in measuring ethological behaviors

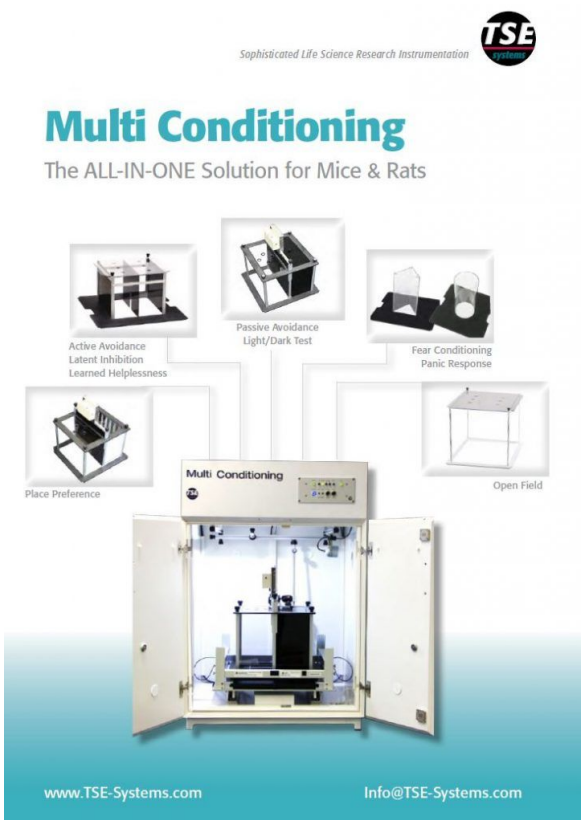
Noldus EthoVision XT14

- video tracking software that tracks and analyzes the behavior, movement, and activity of any animal
- used to acquire all forced swim and elevated plus maze videos and to analyze all of the open field videos
- can also manually or automatically score behavioral events, such as rearing or grooming

noldus.com

TSE Multi Conditioning System

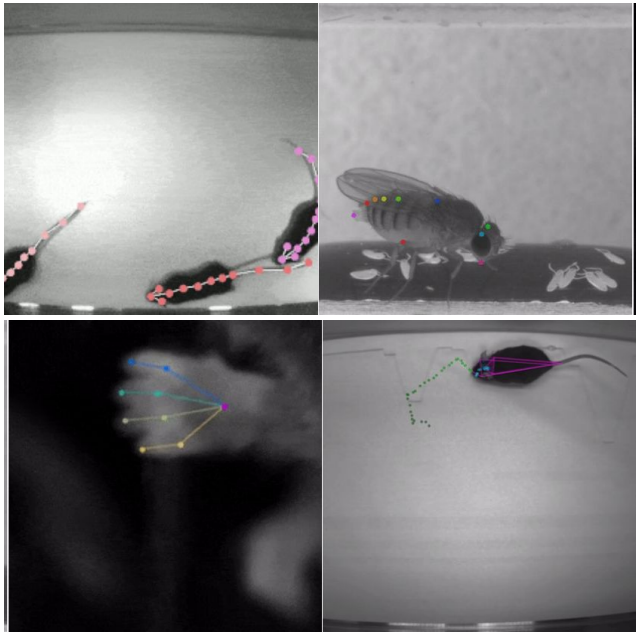
- all in one system that integrates multiple behavioral paradigms for the evaluation of behavior in mice and rats
- uses an infrared beam grid to track movement
- an additional raised beam grid used to measure rearing



tse-systems.com

DeepLabCut

- A software package for animal pose estimation
- efficient method for 2D and 3D markerless pose estimation based on transfer learning with deep neural networks
- allows the user to define and track specific points of interest (e.g. specific body parts)



mackenziemathislab.org

DeepLabCut: labels

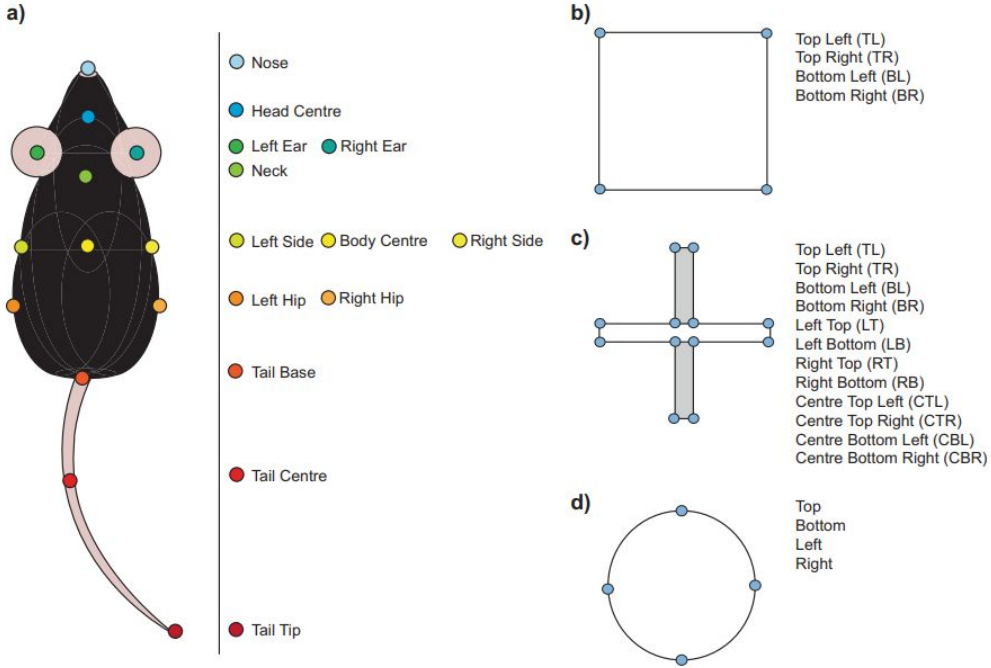
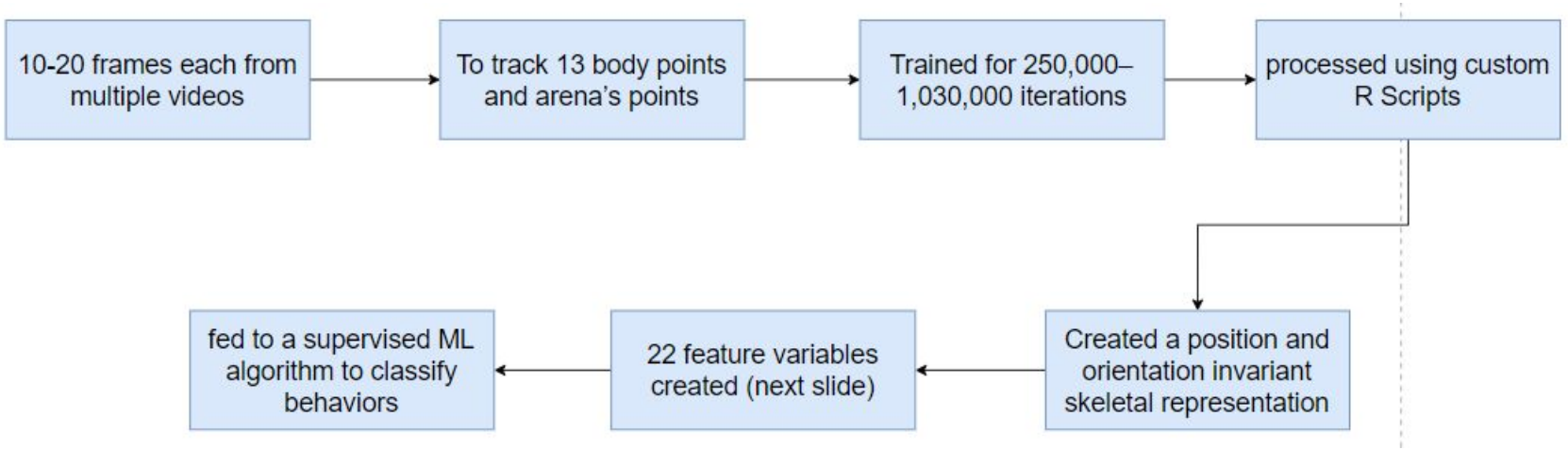


Fig. 1 The labels used to train the DLC networks. a The standardized points of interest used to track the animal. The points of interest required to track the animal in the open field (**b**), the elevated plus maze (**c**) and the forced swim test (**d**).

DeepLabCut: workflow



Original image (Created with app.diagrams.net)

DeepLabCut: feature variables

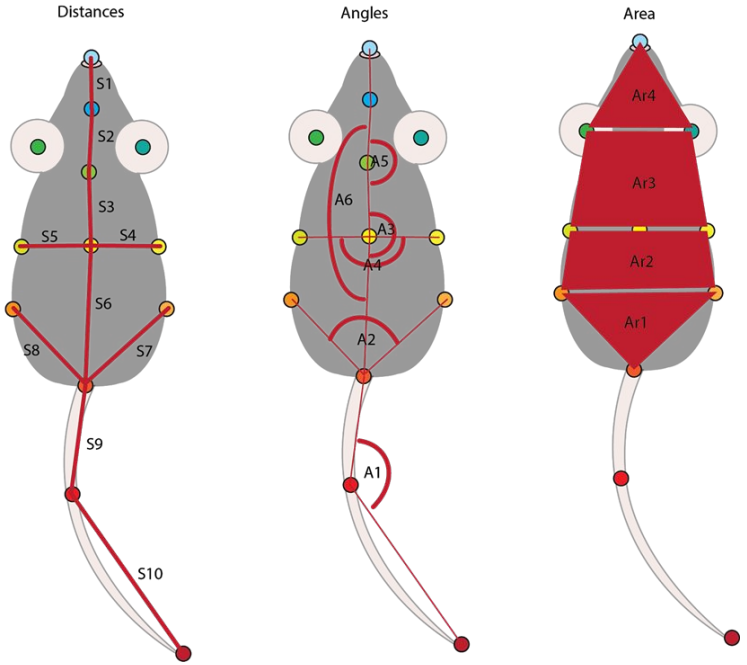


Figure S8. Features used for machine learning. The skeletal information taken into account during the machine learning process, gathered from the tracking data obtained using DeepLabCut.

Results

Tracking Movement

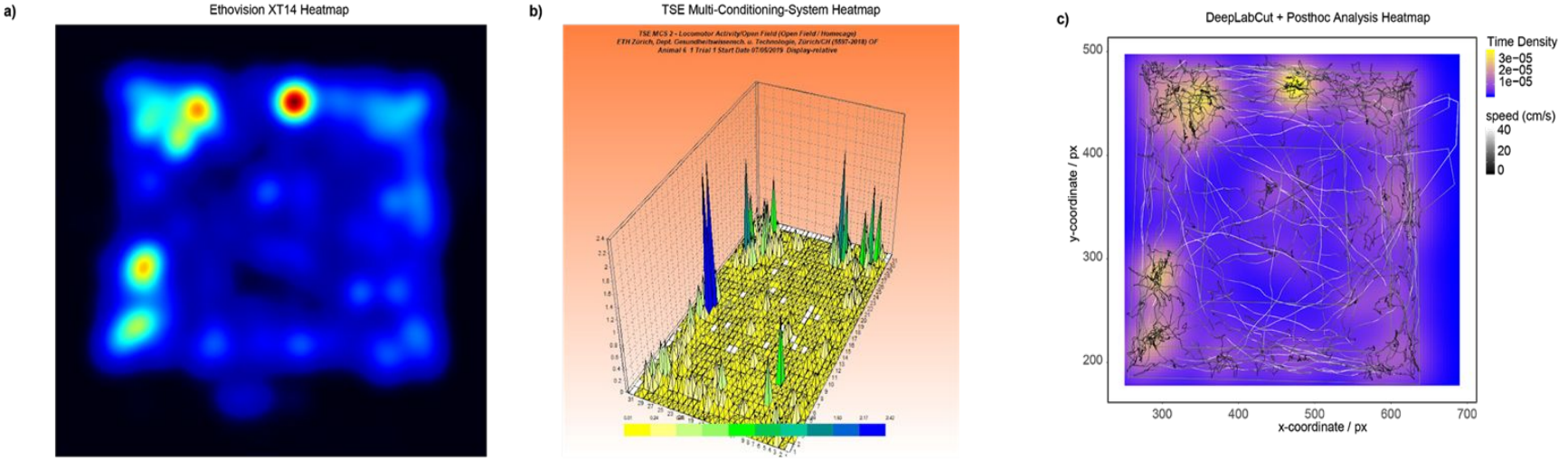


Figure S2. Heatmaps. example heatmaps generated using **(a)** EthoVision XT14, **(b)** TSE's Multi Conditioning System and **(c)** DeepLabCut (with post hoc analysis and ggplot2).

Tracking Movement: Open Field Test

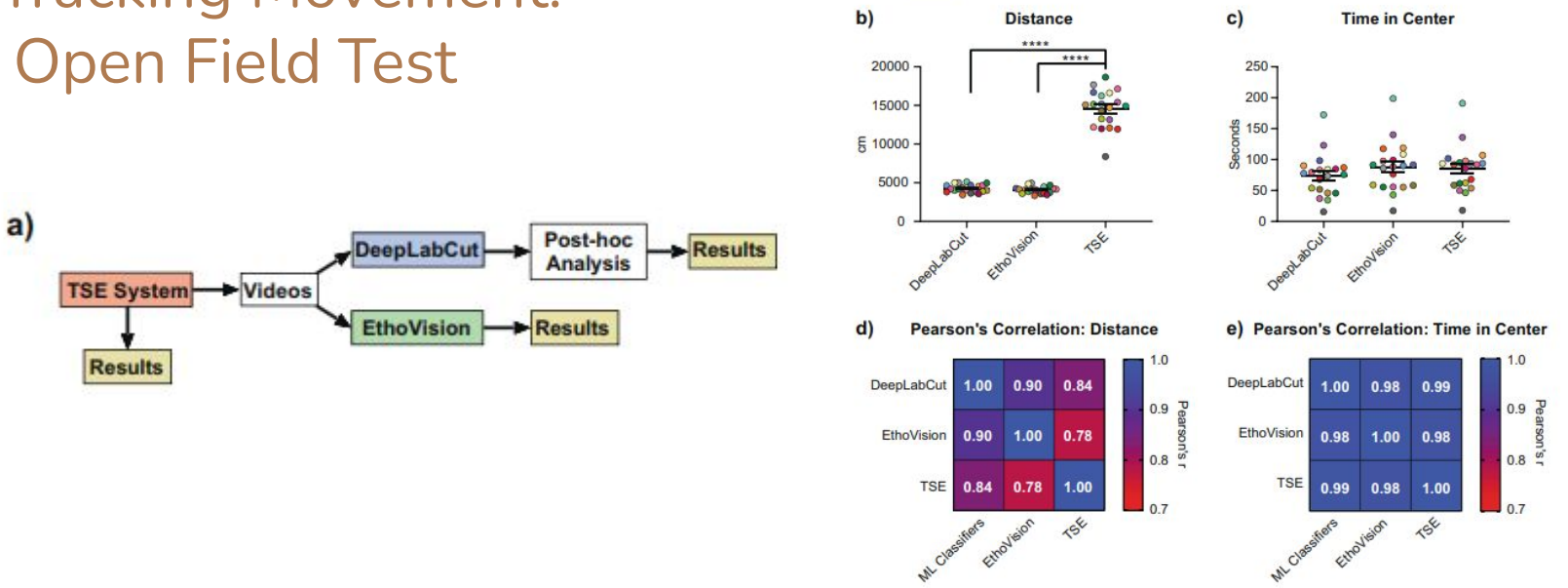


Fig. 2 A comparison of basic tracking parameters in the open field test. a Schematic showing the workflow of the comparison between systems. **b**, **c** Distance and time in center as reported by DeepLabCut (with post-hoc analysis), EthoVision XT14, and the TSE Multi Conditioning System (TSE). **d**, **e** Correlation analysis of the performance of the different systems. Data expressed as mean ± standard error of the mean. Colors represent individual animals and are consistent across analysis techniques for direct comparison ($n = 20$) ****$p < 0.0001$.

Tracking Movement: FST & EPM

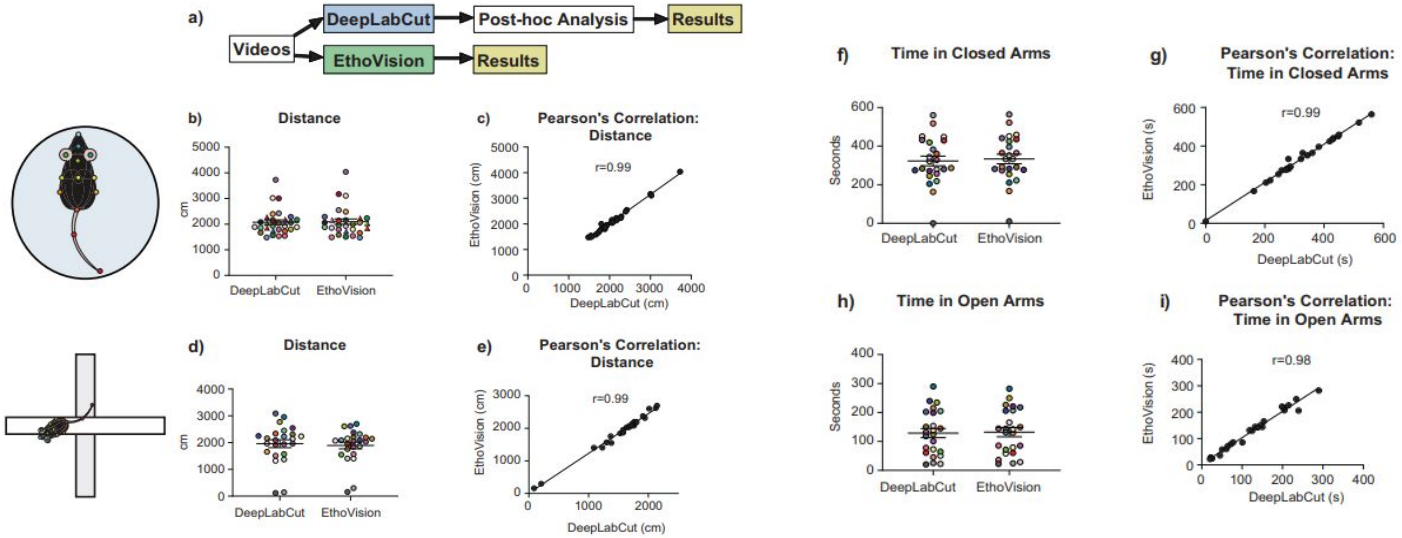


Fig. 3 A comparison of basic tracking parameters in the forced swim test and elevated plus maze. a Schematic showing the workflow of the comparison between systems. **b**, **d**, **f**, **h** Basic tracking parameters in the forced swim test and elevated plus maze as reported by both DeepLabCut (with post-hoc analysis) and EthoVision XT14. **c**, **e**, **g**, **i** Correlation between the scores of the two systems. Data expressed as mean ± standard error of the mean. Colors represent individual animals and are consistent across analysis techniques for comparison (FST $n = 29$, EPM $n = 24$) $^*p < 0.05$.

Ethological behaviors: FST & EPM

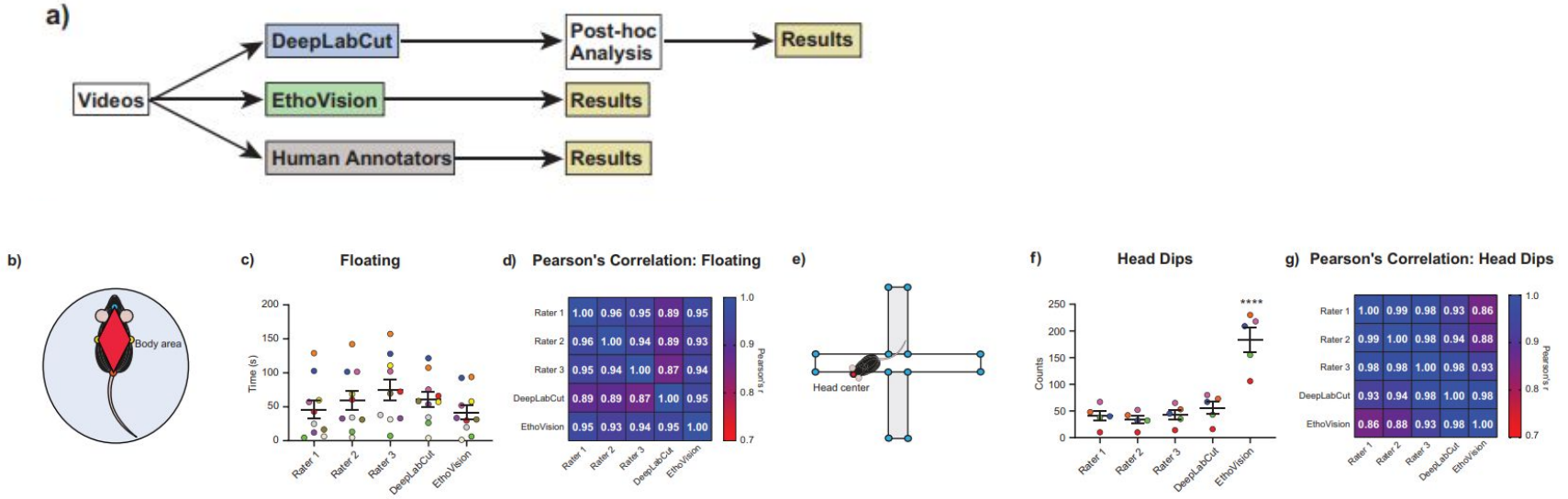


Fig. 4 A comparison of quantifying ethological behaviors in the forced swim test and elevated plus maze. a Schematic of the workflow for the comparison between systems. **b**, **c** The polygon used in the definition of floating, and the body points taken into account when defining head dips. **d**, **e** Floating in the forced swim test and head dips in the elevated plus maze as reported by three human annotators (rater 1–3), DeepLabCut (with post-hoc analysis), and EthoVision XT14's behavioral recognition module. **f**, **g** Correlation analysis for comparison between

Ethological behaviors: Rearing in OFT

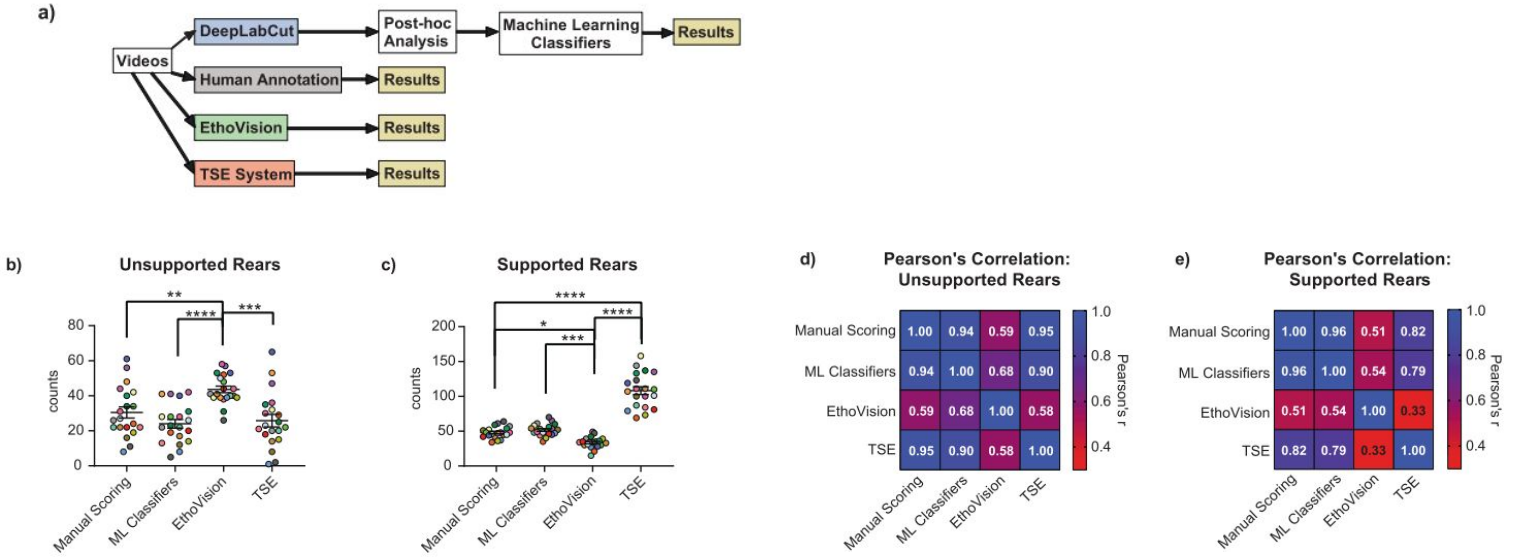


Fig. 5 A comparison of complex behavioral scoring between human raters, machine learning classifiers and commercially available solutions. a Schematic of the workflow. **b**, **c** Unsupported and supported rears in the open field test as reported by three human raters (averaged and plotted as manual scoring) and three machine learning classifiers (averaged and plotted as ML classifiers), EthoVision XT14 and the TSE Multi Conditioning System (TSE). **d**, **e** Correlation analysis for comparison. Data expressed as mean ± standard error of the mean. Colors represent individual animals and are consistent across analysis techniques for comparison ($n = 20$). *$p < 0.05$, **$p > 0.01$, ***$p < 0.001$, ****$p > 0.0001$.

Discussion

Discussion

- DeepLabCut (DLC) video tracking combined with simple post analyses can estimate behavioral data on par or better than commercial solutions
- DLC offers increased tracking flexibility by enabling users to define the parameters of interest themselves
- These approaches achieve human-like accuracy and surpass human reliability, all while being fully automated, flexible, and affordable

Deep Learning for the win!!!

Questions?