Deep Learning imputation of missing data empowers behavior analysis across species

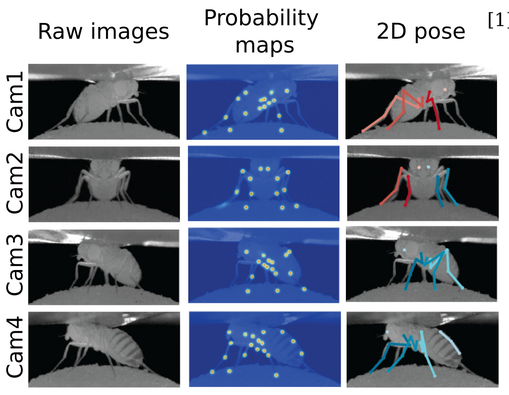
France Rose[1], Monika Michaluk[4], Timon Blindauer[1], Bogna M. Ignatowska-Jankowska[2], Liam O'Shaughnessy[3], Greg J. Stephens[2,3], Talmo D. Pereira[5], Marylka Y. Uusisaari[2], Katarzyna Bozek[1]

1. University Hospital Cologne, 2. Okinawa Institute of Science and Technology, 3. VU University Amsterdam, 4. University of Warsaw, 5. Salk Institute for Biological Studies

Introduction

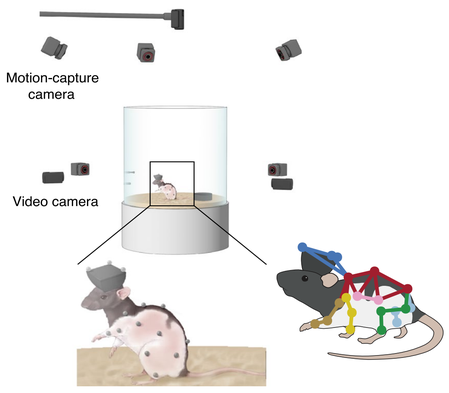
- video pose estimation and motion capture now allow tracking of fine animal movements over extended periods
- these methods can yield low precision detection and missing data
- currently, missing data are dropped or short gaps are linearly interpolated and smoothed (e.g. median or Kalman filter)

Markerless pose estimation



[1]

Motion capture



[2]

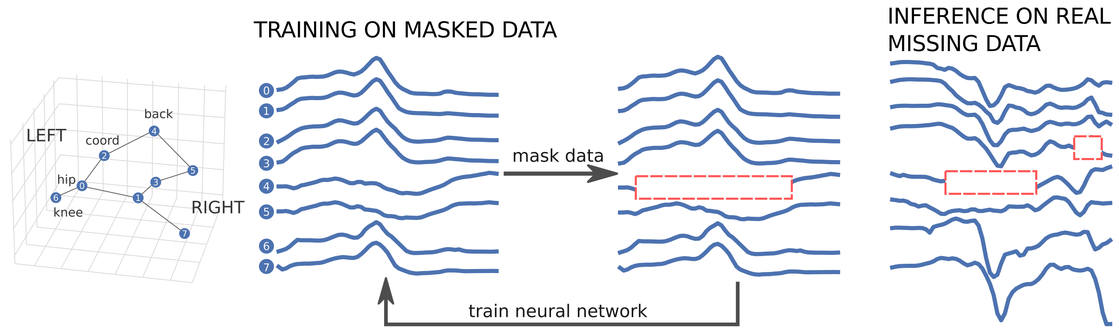
Related work

- **DANNCE**: end-to-end 3D CNN from multiple view videos to 3D pose. Different application than DISK [2]
- **OptiPose**: 3D keypoint refinement (against noise and missing values) [7]
- **BRITS, SAITS, GAN**: general time-series imputation. BRITS inspired our GRU. We tested SAITS' loss without improvement. GANs are known to be hard to train (balance between losses) [8-10]

Transformer-based DISK (Deep Imputation of SKeleton data) imputes accurately missing keypoints in 2D and 3D. A reliable estimated error allows the user to control the quality of the imputed dataset.

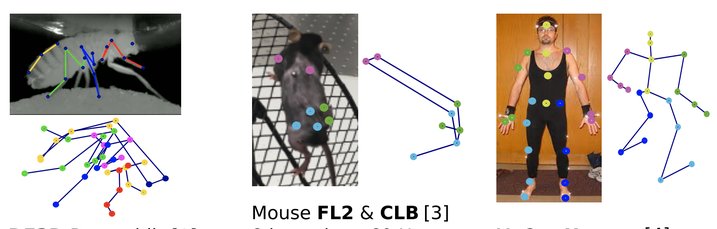
Methods

Self-supervised training, without human annotations



- Estimate the frequency of missing keypoints & gap lengths to train on realistic gaps
- Compare linear interpolation as baseline & deep learning networks:
Gated Recurrent Unit, Temporal Convolutional Network, Graph Convolutional Network, transformer

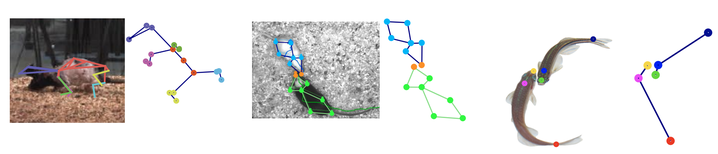
Datasets



DF3D Drosophila [1]
38 keypoints, 100 Hz
0% originally missing

Mouse **FL2** & **CLB** [3]
8 keypoints, 60 Hz
24% & 16%
originally missing

MoCap **Human** [4]
20 keypoints, 12 Hz
0% originally missing

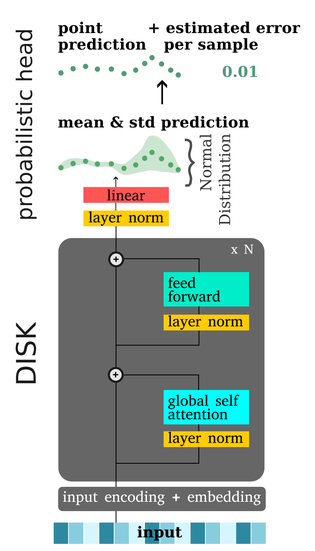


Rat7M [2]
20 keypoints, 30 Hz
44% originally missing

2-Mice-2D [5]
14 keypoints, 30 Hz
0% originally missing

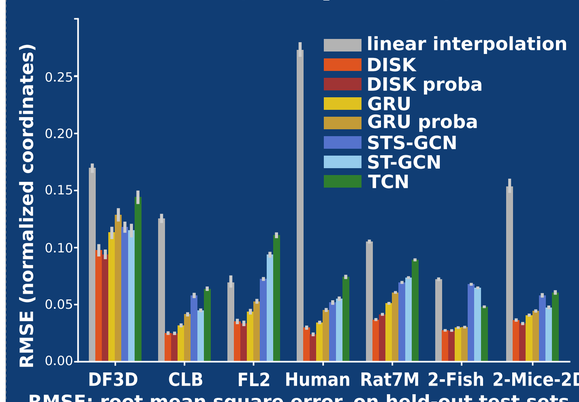
2-Fish [6]
6 keypoints, 60 Hz
4% originally missing

DISK network

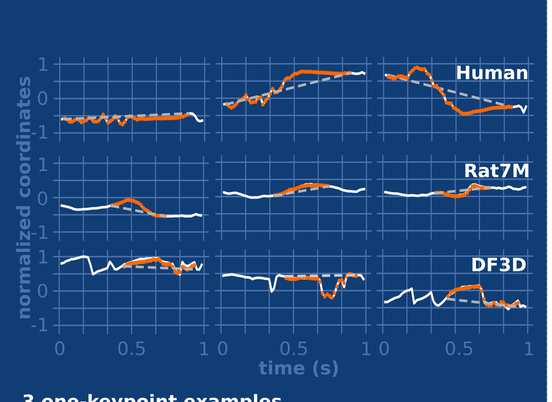


Results

DISK outperforms other methods on 7 datasets (2D and 3D, 5 species, 1 or 2 animals)

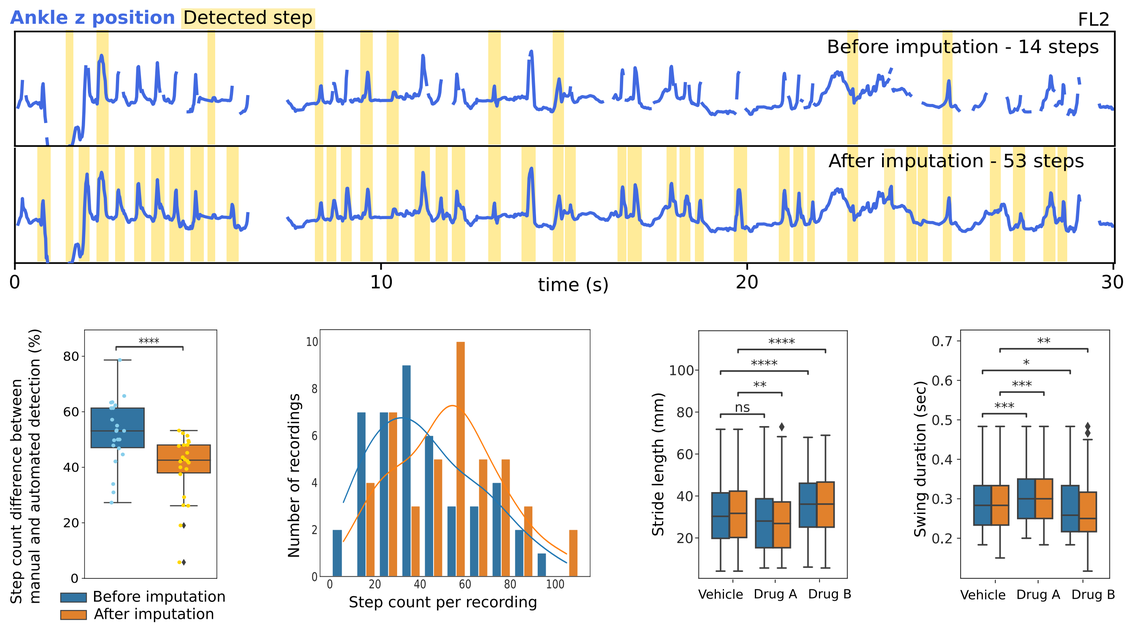


RMSE: root mean square error, on held-out test sets

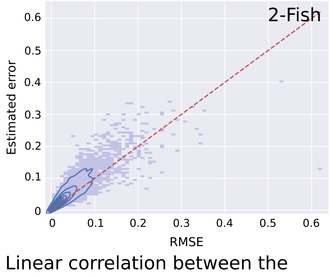


3 one-keypoint examples

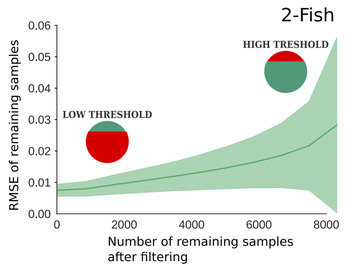
Imputed steps allow finer characterisation of drug effects.



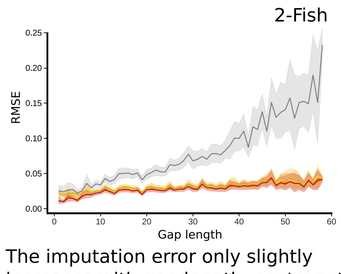
Estimated error allows to filter out below-threshold imputations



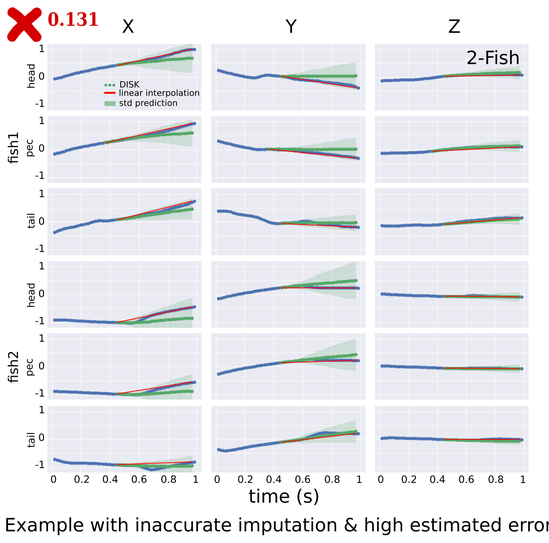
Linear correlation between the estimated and true error.
Pearson correlation: 0.842



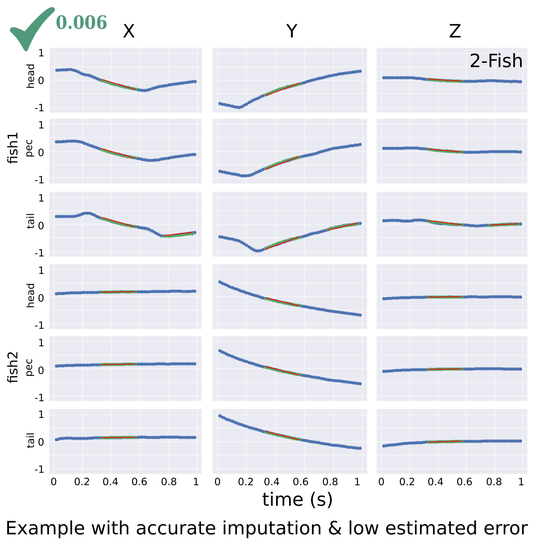
The threshold controls the imputation error in the output data.



The imputation error only slightly increases with gap length, contrary to the linear interpolation.



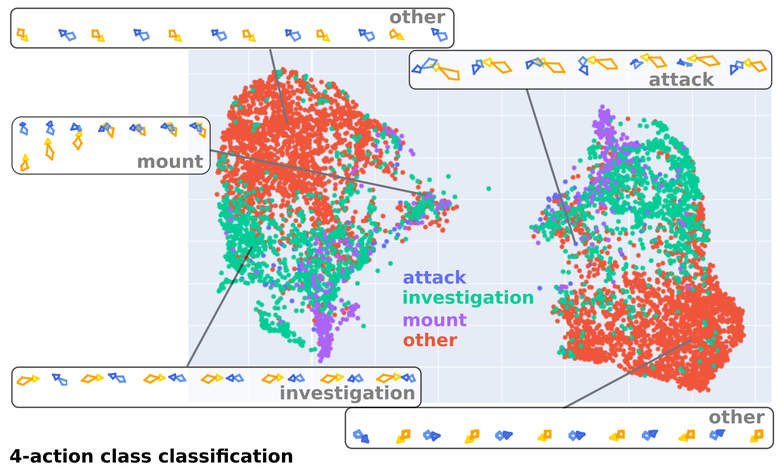
Example with inaccurate imputation & high estimated error

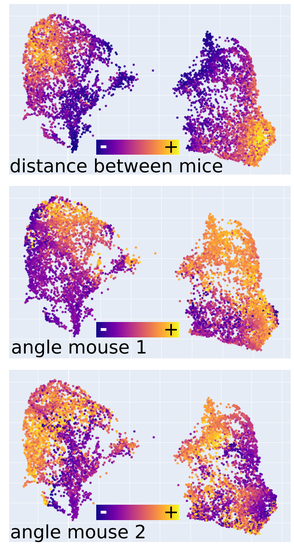


Example with accurate imputation & low estimated error

Learned representations allow to classify behaviors.

U-map on DISK learned representation vectors with action labels — 2-Mice-2D





4-action class classification
with a Random Forest
using the learned representation vectors:
balanced accuracy: 0.877 F1-score: 0.846 precision score: 0.874

[1] Günel et al. eLife 8 (1-23) 2019.
[2] Dunn et al. Nat Met. 18.5 (564-573) 2021
[3] Ignatowska-Jankowska et al. bioRxiv 2023
[4] CMU Graphics Lab Motion Capture Database. 2003
[5] Sun et al. Adv Neural Inf Process Syst (1-15) 2021
[6] O'Shaughnessy et al. bioRxiv 2023
[7] Patel et al. IJCV 131.2 (514-530) 2022
[8] Cao et al. Adv Neural Inf Process Syst 2018.12 (6775-6785) 2018
[9] Du et al. arXiv 2022
[10] Luo et al. Proc Int Joint Conf AI (3094-3100) 2019

UNIKLINIK KÖLN
CMMC CENTER FOR MOLECULARE MEDICINE COLOGNE
BOZEK LAB DATA SCIENCE OF BIOIMAGES
Bundesministerium für Bildung und Forschung
Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen