Using DeepLabCut To Predict Locations of Subdermal Landmarks From Video

Diya Basrai[1], Emanuel Andrada[2], Janina Weber[2], Martin Fischer[2], Matthew Tresch[1,3]

1. Department of Physiology, Northwestern University, Chicago, IL, USA
2. Institute of Zoology and Evolutionary Research, Friedrich-Schiller-University Jena, Germany
3. Shirley Ryan AbilityLab, Chicago, IL, USA

Introduction

- Accurate position tracking of animal subdermal landmarks such as joint centers or skeletal features are necessary for experiments requiring highly accurate kinematic data
- High-speed X-ray acquisition cameras can acquire 3D positions of subdermal landmarks, but manual annotation of positions is laborious.
- Neural-network based markerless tracking software DeepLabCut (DLC)* allows automated position estimation of skin landmarks, but lacks the ability to identify subdermal landmarks
- If we were to train DLC models on X-ray data, could this potentially allow us to predict skin landmarks that correspond to subdermal landmarks?

Here, we present and evaluate an approach that utilizes the automated markerless tracking of DLC to estimate 3D positions of subdermal landmarks from live video

Simultaneous X-ray and live video recording of a rat

- Several trials of a rat running were simultaneously recorded by two live video cameras and two biplanar, high-speed X-ray acquisition cameras
- 9 subdermal landmarks in the rat forelimb, either representing joint centers or skeletal features, were annotated from the X-ray data
- Manual annotation for certain landmarks was helped by prior implantation of tantalum markers

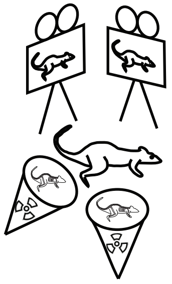
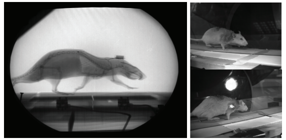
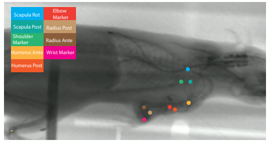


Diagram of camera set-up consisting of two live cameras and two biplanar high-speed X-ray acquisition cameras



Simultaneous views of rat running from X-ray and live video cameras

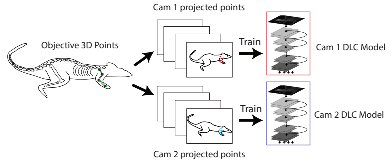


X-ray frame with annotated subdermal landmarks in rat forelimb

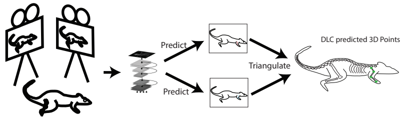
Pipeline to train DeepLabCut models

- 3D positions obtained from X-ray data are projected onto frames from the live video cameras
- DLC models were trained using 80 frames obtained over several trials
- Once trained, the DLC model can predict skin landmarks corresponding to projected positions of subdermal landmarks
- These predicted skin pixel locations can be triangulated to obtain depth-accurate 3D positions

Sample training frame for DLC Cam 1 with projected positions of subdermal landmarks



Pipeline to train DLC models



Pipeline of using trained DLC models to estimate 3D position of subdermal landmarks

Conclusions

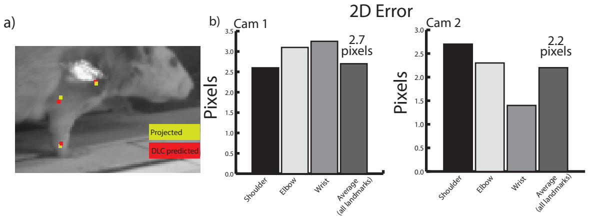
- After an initial training phase that requires X-ray data, trained DLC models were able to accurately predict the locations of subdermal landmarks in the rat forelimb solely from live video
- A pipeline utilizing this approach can be used to speed up the process of obtaining kinematic data and dramatically reduce the amount of labor manually annotating X-ray data.

Future Work

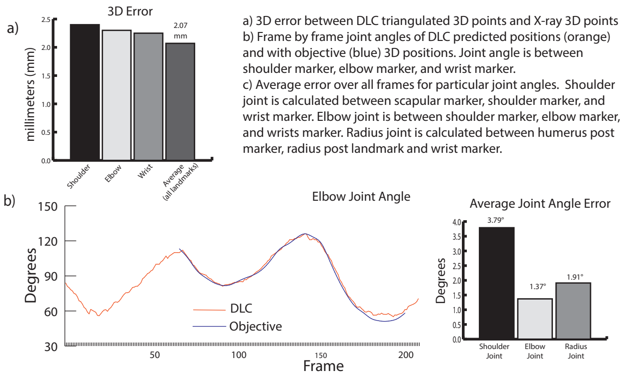
- Immediate future work will focus on applying this approach to predict positions of subdermal landmarks in the rat hindlimb
- Long term goal is to create robust, highly generalizable models. To do so, we are collaborating with various groups that utilize high speed X-ray acquisition on rats
- With a diverse enough set of training data, these trained models would be potentially generalizable enough to predict subdermal locations from across laboratories and experimental conditions

DLC accurately predicts positions of subdermal landmarks in rat forelimb

- Trained DLC models were used to predict 2D positions from a trial neither model were trained on
- Predicted 2D points were compared to projected points obtained from X-ray data
- DLC triangulated 3D points were compared to annotated 3D positions obtained from X-ray data



a) Sample frame (zoomed in) with superimposed predicted and projected labels for shoulder, wrist, and elbow markers
b) Pixel error for DLC models for particular subdermal landmarks. Pixel error is euclidean distance between predicted and projected. Note that shoulder, elbow, and wrist refer to markers representing the joint, not the joint itself.



a) 3D error between DLC triangulated 3D points and X-ray 3D points
b) Frame by frame joint angles of DLC predicted positions (orange) and with objective (blue) 3D positions. Joint angle is between shoulder marker, elbow marker, and wrist marker.
c) Average error over all frames for particular joint angles. Shoulder joint is calculated between scapular marker, shoulder marker, and wrist marker. Elbow joint is between shoulder marker, elbow marker, and wrists marker. Radius joint is calculated between humerus post marker, radius post landmark and wrist marker.

* Mathis et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci, 1281–1289 (2018).

This work is supported by the National Science Foundation under Grant DBI-2015317