

Fig. S1. Training and performance of the DLC Network. A) Error as a function of training iteration for the 15 networks analyzed. Solid lines represent training error and dashed represent test error. Coloring varies from light green to dark blue with increasing training set fractions. Each line shows the mean \pm standard deviation across three shuffles with that training set fraction. (B) Pixel error as a function of training set fraction, evaluated on DLC networks prior to Anipose. Lines represent the mean across shuffles and points indicate individual networks. Train error (solid, light blue) is well below both human labeling errors for all fractions. Test error (dashed, dark blue) drops below inexperienced human labeling error (dot dash, black) at 50% and asymptotes at the experienced human labeling error (solid, black) by the 85% training set. Human labeling errors presented as mean \pm 95% confidence interval. Arrow indicates the network used for further analysis. (C) Median position error (green, solid) and percent frames tracked (orange, dot dash) as a function of training set fraction. This error was measured at the

end of the DLC+Anipose and post-processing pipeline. Median error is not correlated with training set fraction because a reprojection error threshold was applied in post-processing, thus eliminating low-quality frames. Percent of frames tracked, on the other hand, increases with set size because more frames were well-tracked; percent tracked asymptotes at 85%. Arrow indicates the results corresponding to the selected network.

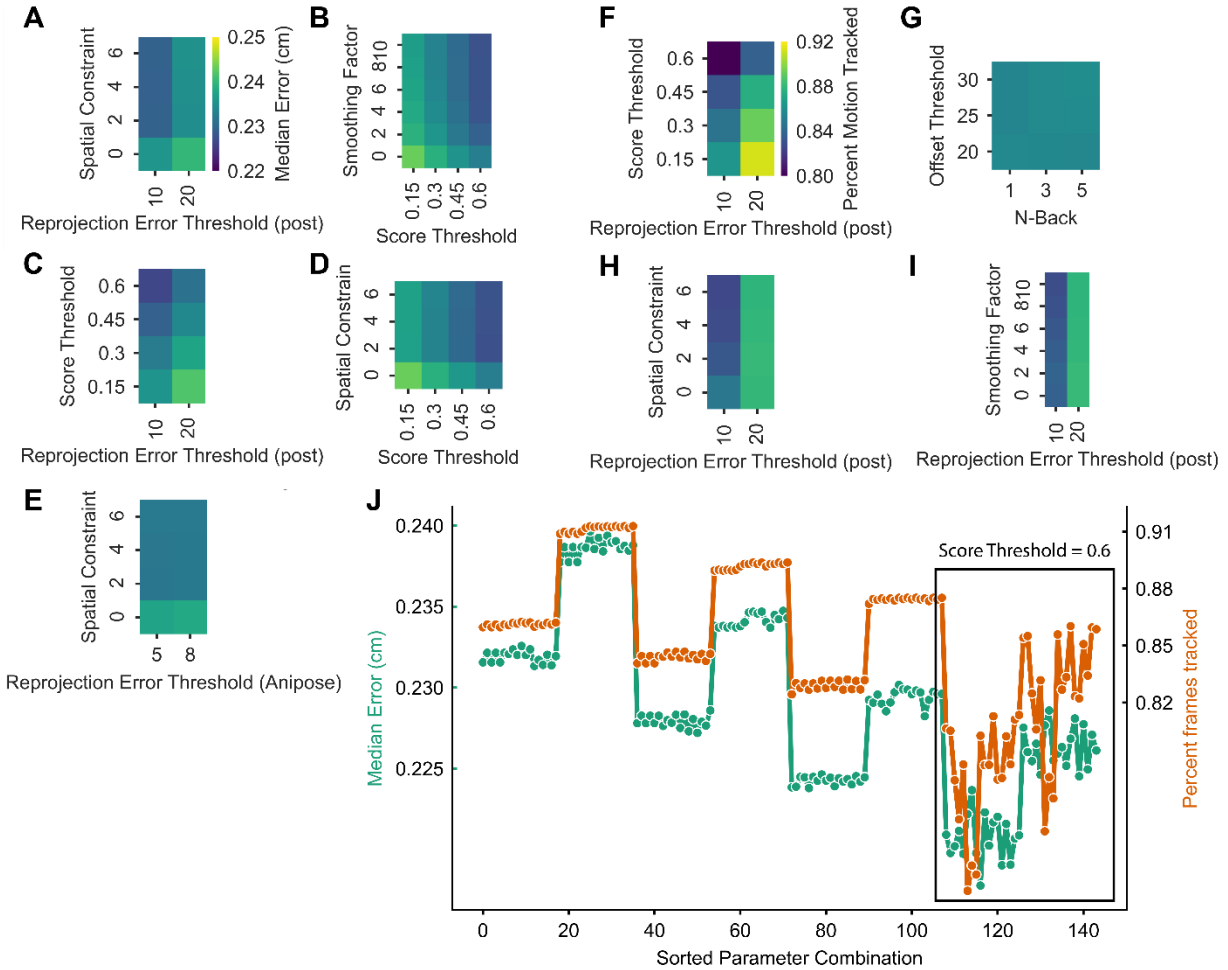


Fig. S2. Interaction between anipose parameters. A) Median error, represented by cell color, versus spatial constraint and post-processing reprojection error threshold. Dark blue cells indicate the lowest error. (B-E) Median error vs four combinations of parameters. (F-I) Similar plots of percent frames tracked vs combinations of parameters, where yellow cells correspond to the highest percent and blue cells to the lowest. Note that the interactions in (A-I), although technically identified as significant, are minor and not evident given the vertical and horizontal striations in the figures. (J) Median error and percent tracked for 144 parameter sets including all score threshold values with smoothing factor and spatial constraint set to 6 and 2, respectively. Results are sorted along the x-axis by score threshold first, post-processing reprojection error threshold second and n-back third. The median error and percent tracked vary predictably for score thresholds between 0.15 and 0.45, but combinations with score threshold = 0.6 vary unpredictably with the other parameters.

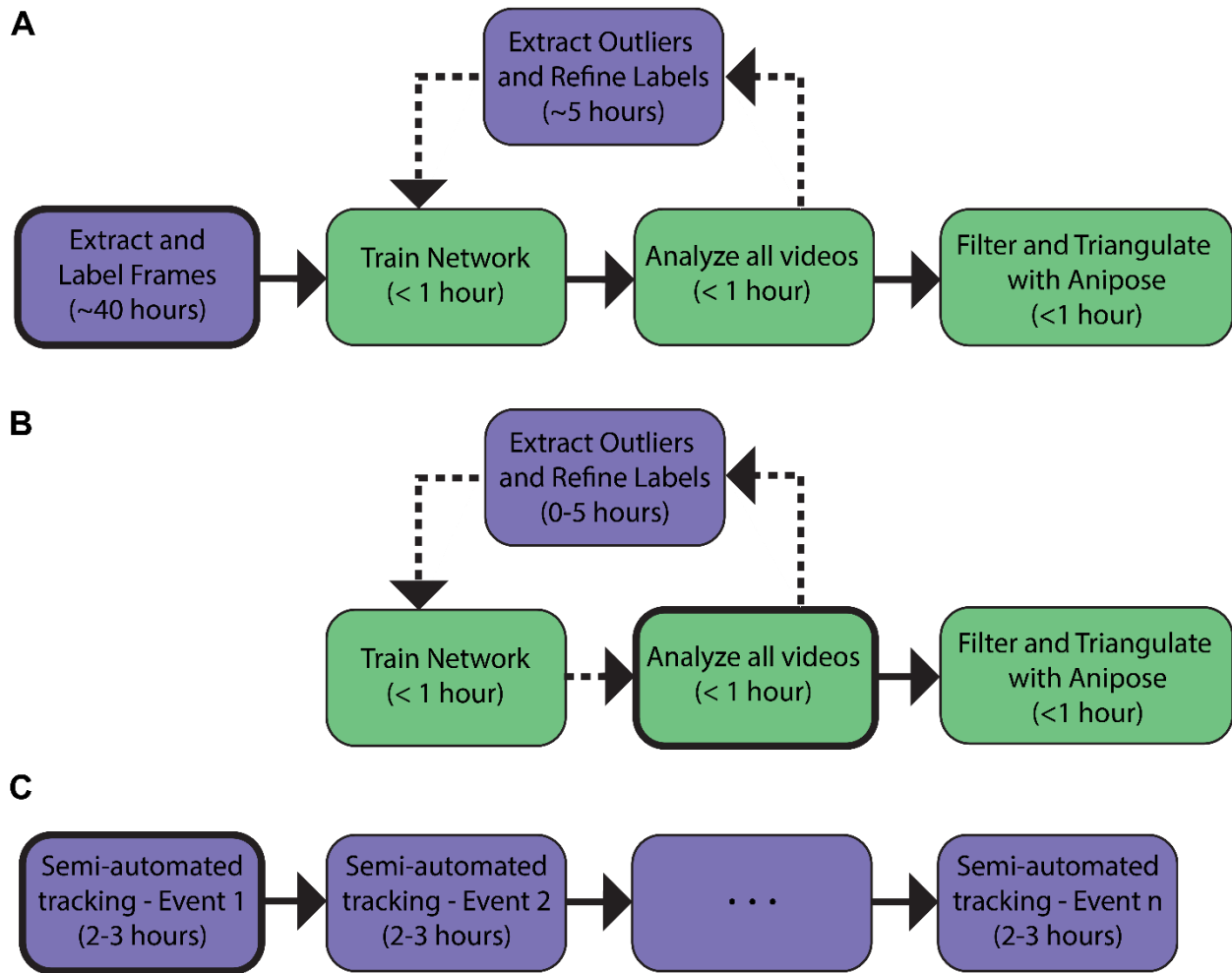


Fig. S3. DeepLabCut and XROMM pipelines. A) The DLC workflow for the first time working with video data in a novel setting. B) DLC workflow for analysis of subsequent data with the well-trained network from (A). (C) The standard XROMM workflow. Bold outlines indicate the starting point of each process. Purple steps require significant hands-on work while green steps are primarily computational. All time estimates provided refer to hands-on work either completing manual steps or preparing for computational steps. Dotted lines indicate an optional path in the pipeline. Calibration steps are not very different between DLC and XROMM and are not shown.

Table S1. Corresponding XROMM and DLC target locations. XROMM markers were targeted superficial to specific skeletal or muscular locations. DLC labels were chosen to match XROMM locations as close as possible, although we had to adjust marker positions to the location of clear visual landmarks so that consistent human labeling was possible. Colors match the colors in Fig. 1

XROMM Location	DLC Location	Color
Body of metacarpal 3	Base of metacarpal 3	●
Base of metacarpal 4	Proximal to base of metacarpal 3	●
Base of metacarpal 2	Base of metacarpal 2	●
Flexor carpi ulnaris	Visual landmark on distal forearm	●
Brachioradialis	Visual landmark on medial forearm	●
Anconeus	Elbow	●
Lateral tricep distal	Visual landmark on distal upper arm	●
Lateral tricep proximal	Visual landmark on medial upper arm	●
Deltoid	Visual landmark on proximal upper arm	●
Vertebrae – T4		●
Vertebrae – T8	Visual landmark on torso (posterior)	●