

Heimdallr: A Dataset for Sport Analysis

Michael Riegler, Simula & UiO
Duc-Tien Dang-Nguyen, University of Trento
Bård Winther, Simula
Carsten Griwodz, Simula & UiO
Konstantin Pogorelov, Simula & UiO
Pål Halvorsen, Simula & UiO





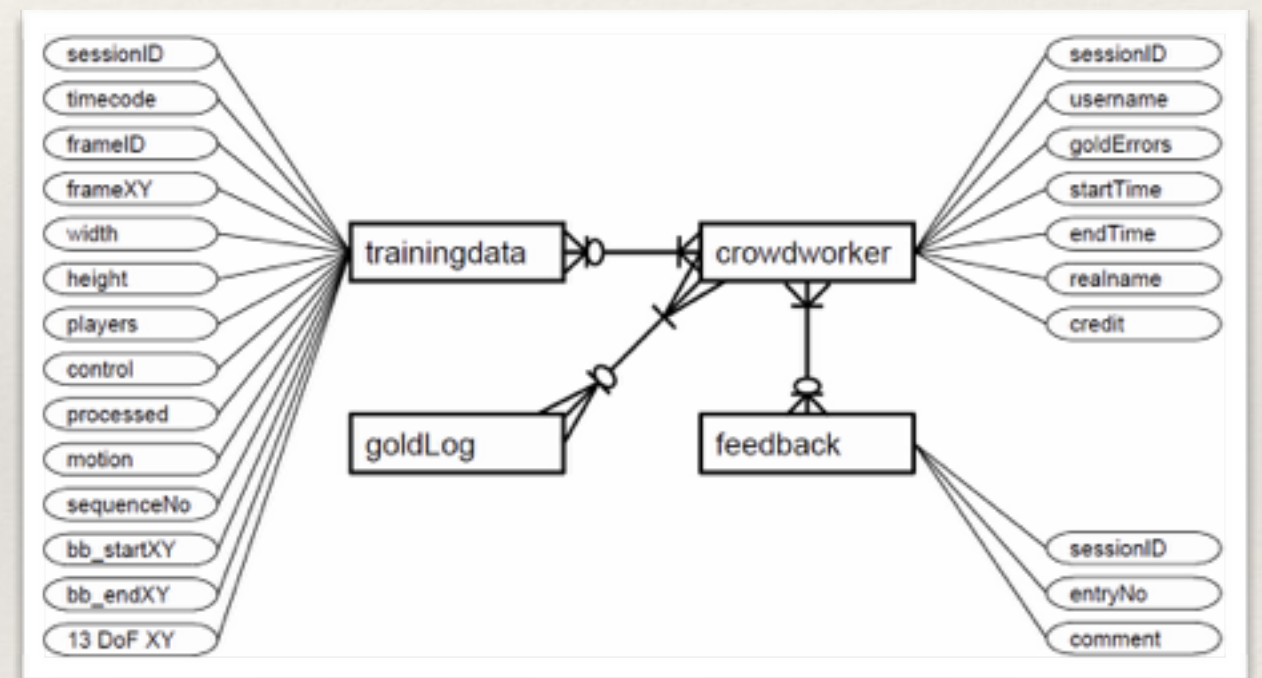
The Idea

- ❖ Collect a dataset of annotated soccer scenes
- ❖ Two purposes
 - ❖ Action recognition and pose estimation
 - ❖ Improved understanding of crowdsourcing workers
- ❖ Apart from that, we provide the application used for collecting data



The Dataset

- ❖ More than 3.000 fully annotated frames
- ❖ 42 different sequences
- ❖ Over 10.000 written feedback
- ❖ 592 different workers
- ❖ Useful for researchers looking into pose estimation and crowdsourcing



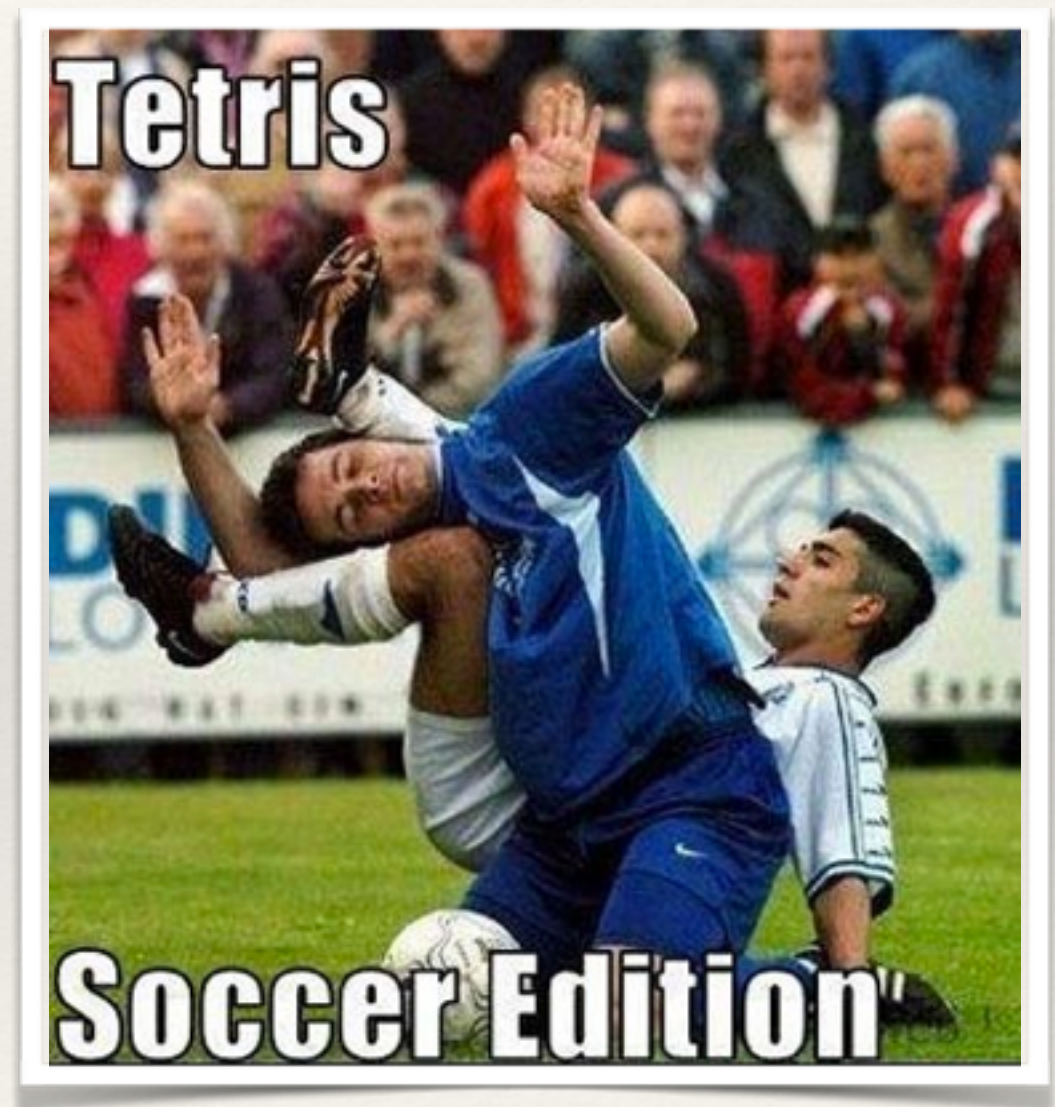
Differences to Existing Datasets

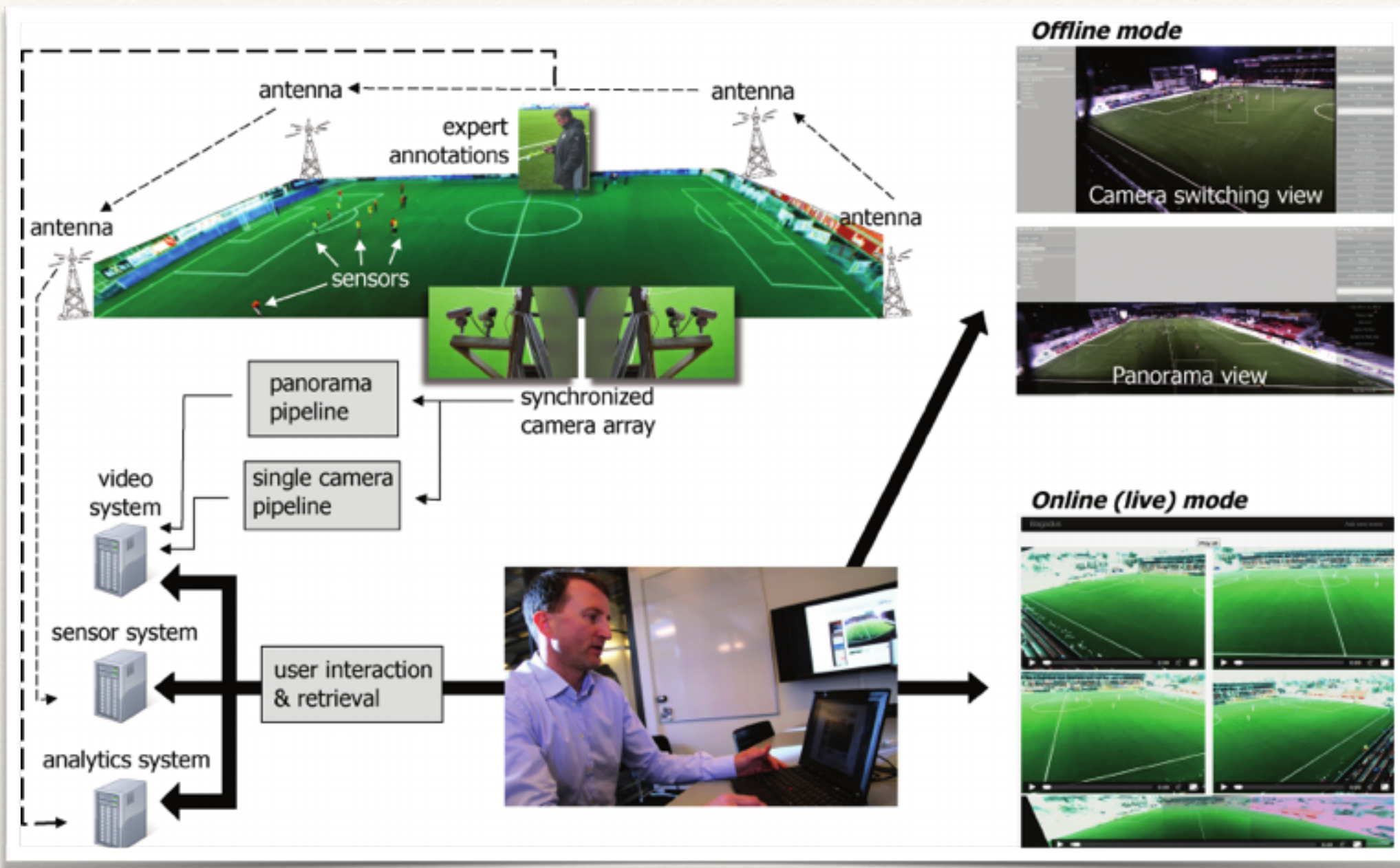
- ❖ Not only close-up shots of players, but also...
 - ❖ External calibration of the camera with respect to the field
 - ❖ x and y positions of the players
- ❖ All scenes are taken by one static camera-array system
- ❖ All collected crowdsourcing data and our filtering as a possible ground truth



Data Collection

- ❖ 3 main steps
- ❖ Scenes collected using the Bagadus system
- ❖ Crowdsourcing to collect user annotations
- ❖ Quality and filtering methods for the crowdsourced data





Sequences

- ❖ 42 different sequences
- ❖ Run, sprint, walk, walk-backwards, side-jump, kick
- ❖ Consisting of 18 to 168 frames

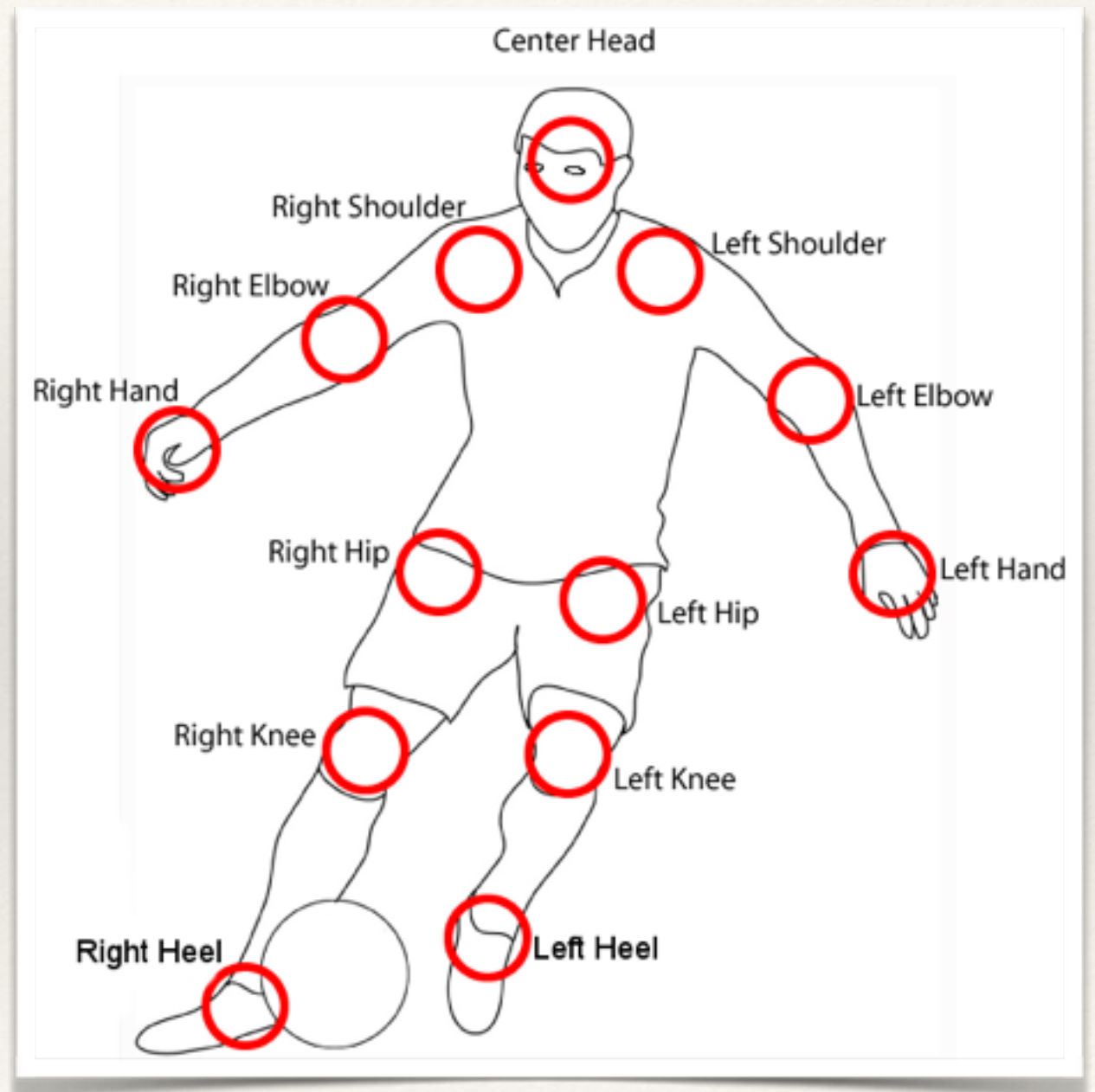
Crowdsourcing

- ❖ Performed using Microworkers
- ❖ 592 different workers
- ❖ Experts annotations as ground truth (people that are experienced with soccer and the data)
- ❖ One worker annotated ca. 48 frames per hour



Annotations

- ❖ 13 joints of the human body
 - ❖ Head, shoulders, elbows, hands, hips, knees and feet
- ❖ Using a online training tool
- ❖ Frames are randomly assigned to workers
- ❖ Motion label (which action was performed)



Online Training Tool

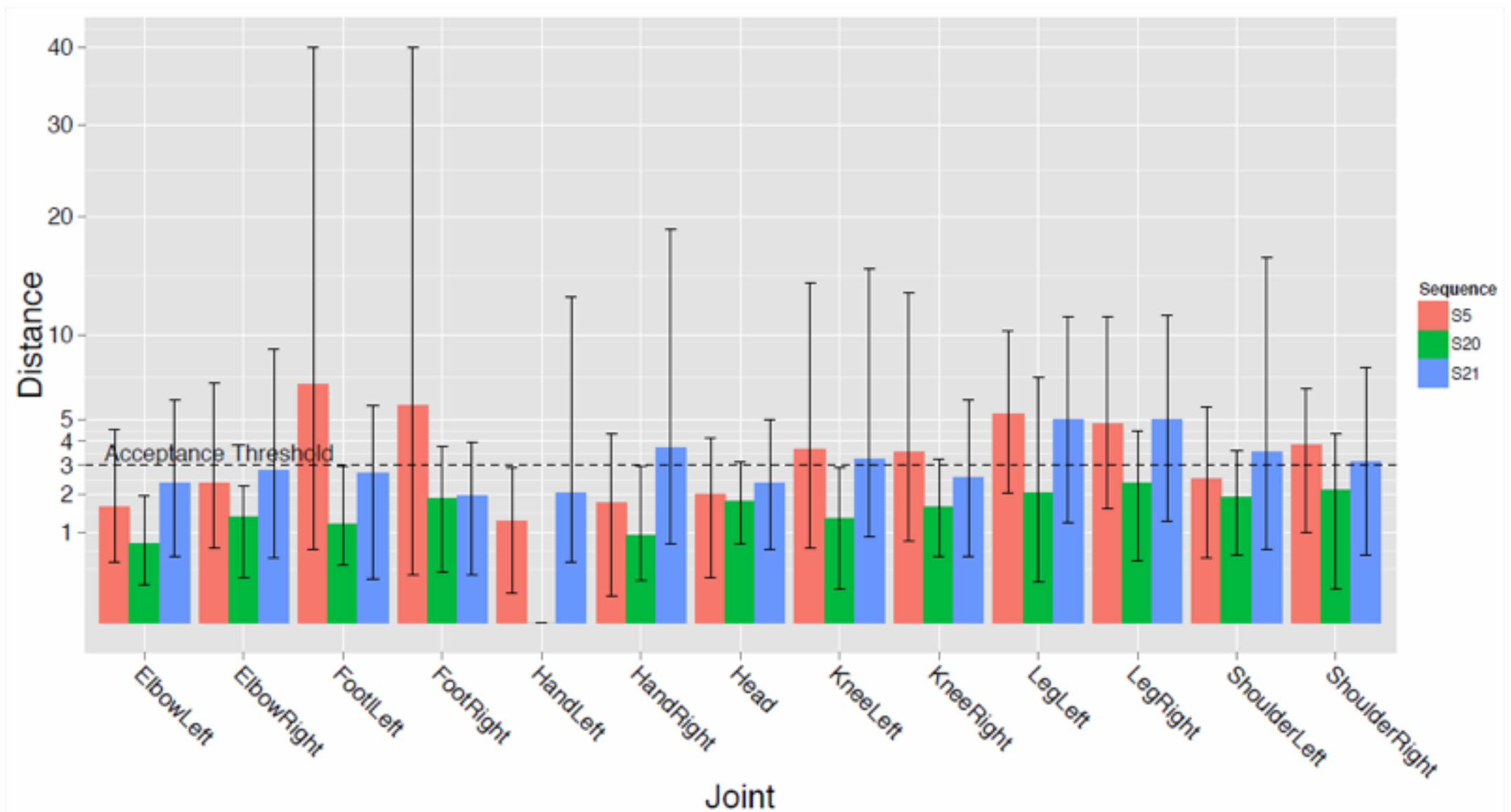
Apply Selection

New Image Reset Undo No Player Obscured Commit <Motion> 1 Player



Please click on the **Left Heel**

0:02



Crowdworkers Performance

Annotation performance of crowdworkers for 3 sequences

Applications of the Dataset

- ❖ Action classification
- ❖ Pose estimation
- ❖ Crowdsourcing quality
- ❖ Workers quality
- ❖ Outlier detection
- ❖ Many more...



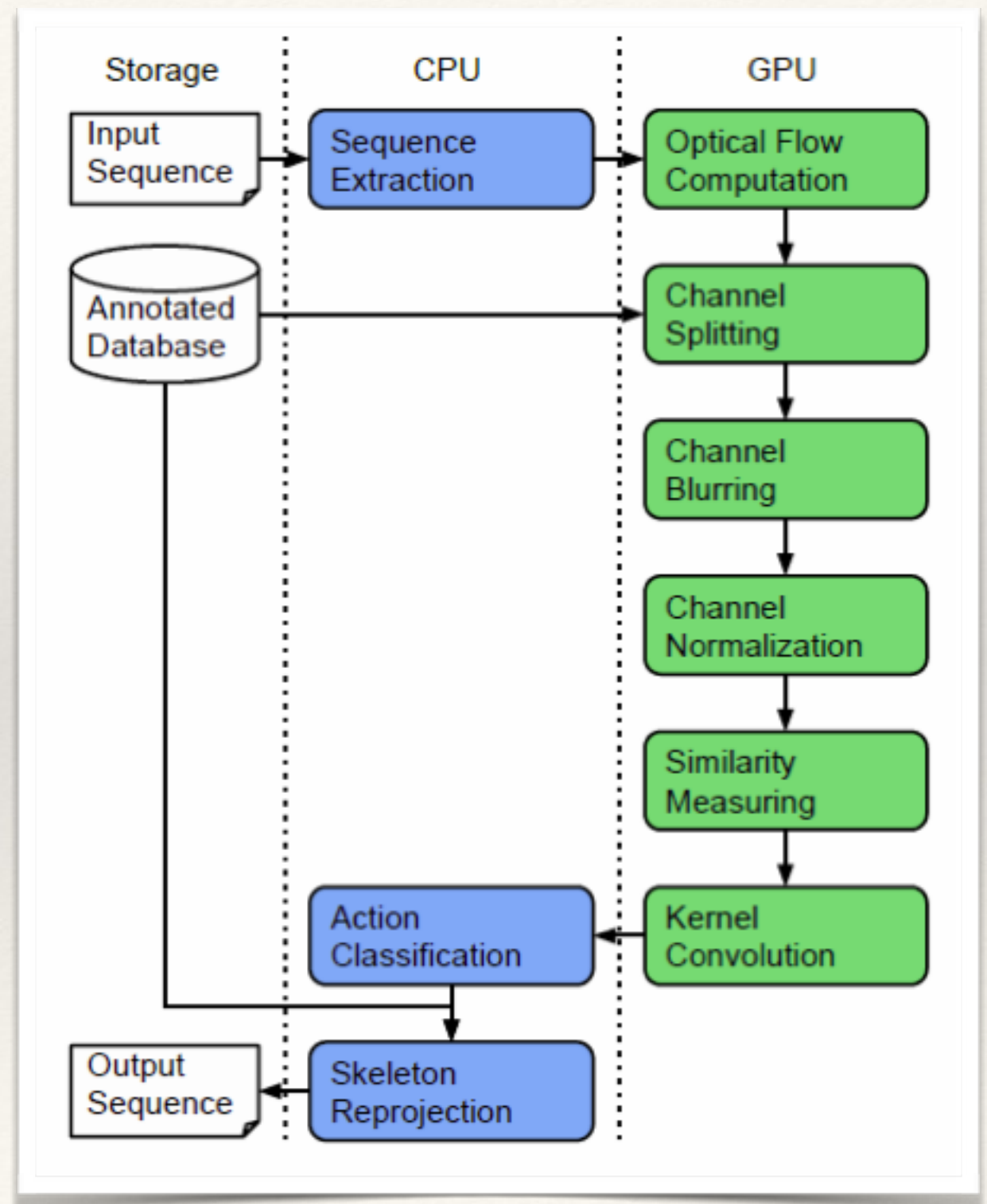
Crowdsourcing Quality Control

- ❖ Finding workers who try to cheat
- ❖ 3 main ways of cheating identified
 - ❖ Cluster, lines and random
- ❖ By filtering and using majority vote we could obtain good skeletons



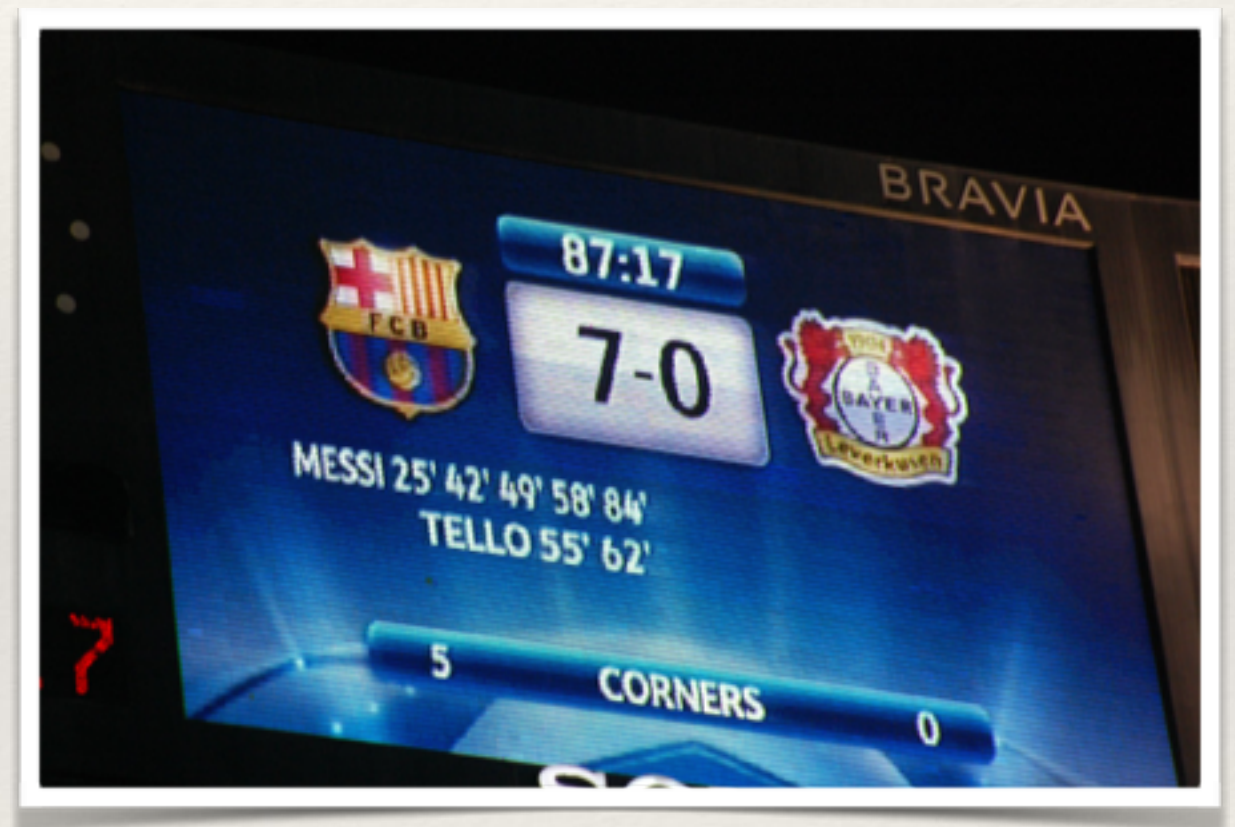
Action Classification

- ❖ Simple Nearest Neighbour algorithm
- ❖ Around 75% of all sequences correctly classified
- ❖ Up to pixel perfect poses were estimated
- ❖ Can also be considered as a baseline for users of Heimdallr



Summary

- ❖ Heimdallr can be an interesting dataset for two groups of researchers
- ❖ Allows to address different tasks such as action classification, pose estimation, worker discarding, workers quality estimation, etc.
- ❖ Training tool is provided with the dataset as open source software



Thank You and Questions?



[**simula**]



UNIVERSITY OF TRENTO - Italy

CrowdMM 2017



@MMM 2017 in Iceland