Protocol 13
Data Preparation

05.04.2019

In order to use the Spatial Temporal Graph Convolutional Networks (ST-GCN), the input data should be converted to .npy (data) and .pkl (label) files.

1. Skeleton (Keypoints) to JSON files

The authors of ST-GCN extracted human skeletons from videos by using Openpose. In this project, Deeplabcut is used to extract the skeletons of cows, and the file "utils/make_labeled_video.py" is modified to output .json files for each video. There are some issues regarding this step:

- The parameter "pcutoff" should be set to 0 to ensure that each frame contains all the keypoints, even though the prediction may not be accurate. The reason is that ST-GCN requires that each frame has exactly the same number of keypoints to construct a graph.
- The number of frames for each file should be less than 300 according to ST-GCN.
- The input data of ST-GCN have class labels. For the lameness detection as a regression problem, the locomotion scores (LS) are divided into four intervals, each of which is a class:
  - Class1: LS 1-2
  - Class2: LS 2-3
  - Class3: LS 3-4
  - Class4: LS 4-5 (containing 5)

  This categorization will be modified, such that the output from the network is a floating point rather than a class.

2. JSON to .npy and .pkl files

The .json files for each video need to be split into training and test sets ("data_split_st-gcn.ipynb"), as well as to be converted to .npy (data) and .pkl (label) files. Two files from st-gcn (https://github.com/yysijie/st-gcn) were used and modified:

- "feeder/feeder_kinetics.py"
- "tools/kinetics_gendata.py"