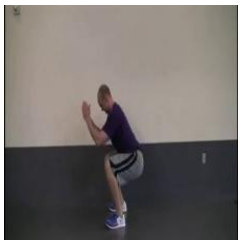
# Introduction to Deep Learning in Computer Vision
# Project 4.1: Video classification

In this project you will be classifying videos according to their actions. You will work on a subset of the ufc-101 dataset for video action recognition derived from this Kaggle version of UFC-101. You can find the designated dataset as well as the file `datasets.py` on Learn.

## The dataset and provided tools

The provided dataset includes 720 videos (500/120/120 for train/val/test) of 10 balanced classes related to workout, mode precisely including the classes ['BodyWeightSquats', 'HandstandPushups', 'HandstandWalking', 'JumpingJack', 'JumpRope', 'Lunges', 'PullUps', 'PushUps', 'TrampolineJumping', 'WallPushups']


BodyWeightSquats


HandstandPushups


HandstandWalkin


JumpingJack

JumpRope


Lunges


PullUps


PushUps


TrampolineJumping


WallPushups

For each video, we provide:

- 10 uniformly sampled frames
- csv files which include the video name, relative path, actions, and labels.

Moreover, we provide example dataset classes and dataloaders in `datasets.py` to help you get started with both 2D or 3D modeling. You may find the dataset of `frames` useful for training per-frame models and the `videos` dataset useful either for testing on videos or training early/late fusion models (if stack_frames=True returns the sampled frames are stacked as [C, T, H, W], else as a list).

# Your task

In this first part of the project, we your task is to explore different models for video classification, namely:

- aggregation of per-frame models
- late fusion
- early fusion

… all for 2D CNN models, along with

- 3D CNNs.

Please report your comparative results in your poster. To work with the models – in particular, the 3D CNN – you may have to subsample the images. Please include any such details in the poster and discuss the potential effects thereof.

Next week, you will receive additional tasks based on the associated lecture.