Deep Learning Winter School for Computer Vision 2020

Practical

11 February 2020

Pre-Requisite-

- Download the data - 1. **3D poses** [Link] and 2. **RGB data** [Link]
- Save the 3D poses on your google drive and unzip them. This data will be used (by mounting your gdrive) to execute skeleton based Action recognition using LSTM.
- For the RGB data, the pre-processed human bounded crops are provided in a zip file. Unzip it and uniformly sample 30 frames out of each video. Save them on your google drive. This data will be used for RGB based Action Recognition using Recurrent Convolutional Network.

Dataset Description - MSRDailyActivity3D - comprises of 320 videos. There are 16 Actions performed by 10 subjects in two different environment. For this practical, we will follow one-subject-out protocol. The subject with identifier 's08' will be in the test set and all other subjects will be in the training set.

1. **Skeleton based Action Recognition** - Implement a 3 layer stacked LSTM for skeleton based Action Recognition. This framework will be illustrated in the class (as an example). Use the same git code (mentioned in the example) to solve the next problem. [Link to google colab]
2. **Recurrent Convolutional Network for Action Recognition** - Extract the 2D convolutional features from your favourite 2D CNN for every sampled frames. Then feed these frame-level features to LSTM for Action Recognition.

Note the difference between both the Action Recognition frameworks.