

Datasets Preparation

Pravin

October 2019

1 CNN features:

For CNN features we sampled the video at 2.14fps and used ‘GoogleNet’ which gives 1024 dimension feature vector for a single video frame. For example the duration of ‘Alin_Day_1’ is 5 hrs 39 mins 26 secs (after combining all small clips as suggested in paper). So we get 43643×1024 feature vector.

You can use following code to extract CNN features for the uploaded h5 file:

```
import h5py
h5py.File('Disney_features_CNN_C3D.h5','r')
feat = f['Alin_Day_1.mp4']['features']
```

2 C3D features:

We extract C3D features for each second i.e we get 512 dimension feature vector corresponding to each second of the video. So we get 20365×1024 You can use following code to extract C3D features for the uploaded h5 file:

```
import h5py
h5py.File('Disney_features_CNN_C3D.h5','r')
feat = f['Alin_Day_1.mp4']['features_c3d']
```