Our Codes are summarized as follows:

- **Feature extract**. Models we extract features:

Motion(3d resnet): https://github.com/kenshohara/3D-ResNets-PyTorch

Object(Faster rcnn): https://github.com/endernewton/tf-faster-rcnn

Scene: (Places205-VGGNet):https://github.com/wanglimin/Places205-VGGNet

Pose(Openpose):

https://github.com/ZheC/Realtime_Multi-Person_Pose_Estimation

After extract features from original videos, we place them in the 'caches' folder.

- **Train**.

1、Compute the knowledge map with the file 'Compute_mean_vector'.

2、Divide the video classes into groups use the file 'Cluster_to_divide_group'.

3、Separately train the coarse model and fine models use files 'Classifier_coarse' and 'Classifier_fine'.

- **Test**.

Fusion:Use the file 'Classifier_fusion' to train the features fusion without coarse-to-fine model then test.

Coarse-to-Fine:Use the file 'Coarse_to_fine_model ' to test our features fusion with coarse-to-fine model.