

## Video+Audio+Text-based Multimedia Event Detection

Name: Chethan Singh Mysore Jagadeesh

AndrewID: cmysorej

### Problem Description:

Perform multimedia event detection(MED) with video, audio and text features.

### Features and results:

#### Early Fusion: two experiments conducted:

1. 100 features from Mobilenet V2 features + 50 features from Soundnet16 Kmeans BoWs + ASR features
2. 100 features from Mobilenet V2 features + 50 features from Soundnet16 GMM BoWs + ASR features

#### Double Fusion:

1. Got Scores for individually for Mobilenet V2 features for 3 events seperately
2. Got Scores for individually for 50 features from Soundnet16 Kmeans BoWs for 3 events seperately
3. Got Scores for individually for ASR for 3 events seperately
4. Got Scores for individually for (Mobilenet V2 features + 50 features from Soundnet16 Kmeans BoWs) for 3 events seperately
5. Got Scores for individually for (Mobilenet V2 features + ASR for 3 events) seperately
6. Combined 5 scores for each event seperately. So after this, I got final score file each for an event and used this as features to train the final SVM classifier

	Features	P001	P002	P003
Early Fusion 1	MobilenetV2 + soundnet16_kmeans + ASR	Average precision: 0.762814052273814 Kernel: rbf Regularization Param(C): 10.0 Gamma: scale	Average precision: 0.84463630099465 Kernel: rbf Regularization Param(C): 0.01 Gamma: scale	Average precision: 0.6165516274394861 Kernel: linear Regularization Param(C): 80.0 Gamma: scale
Early Fusion 2	MobilenetV2 + soundnet16_gmm + ASR	Average precision: 0.7778564386979785 Kernel: sigmoid Regularization Param(C): 0.1 Gamma: scale	Average precision: 0.8691891860044819 Kernel: rbf Regularization Param(C): 5.0 Gamma: scale	Average precision: 0.627707563643102 Kernel: rbf Regularization Param(C): 5.0 Gamma: auto

<b>Double fusion</b>	<b>MobilenetV2 score + soundnet16_gmm score + ASR score + (MobilenetV2 + soundnet16_gmm) score + (MobilenetV2 + ASR) score</b>	Average precision: 0.2684245543185012 Kernel: rbf Regularization Param(C): 0.01 Gamma: auto	Average precision: 0.42859924155162366 Kernel: rbf Regularization Param(C): 0.03 Gamma: auto	Average precision: 0.3757365913487467 Kernel: rbf Regularization Param(C): 60.0 Gamma: scale
----------------------	--	--	---	---