Input Video Clip

Convolution 3D : 128

MaxPool 3D

Convolution 3D : 256

MaxPool 3D

Convolution 3D : 75

MaxPool 3D

Flatten

BiLSTM 128×2

BiLSTM 128×2

Dense