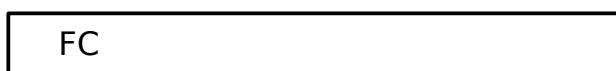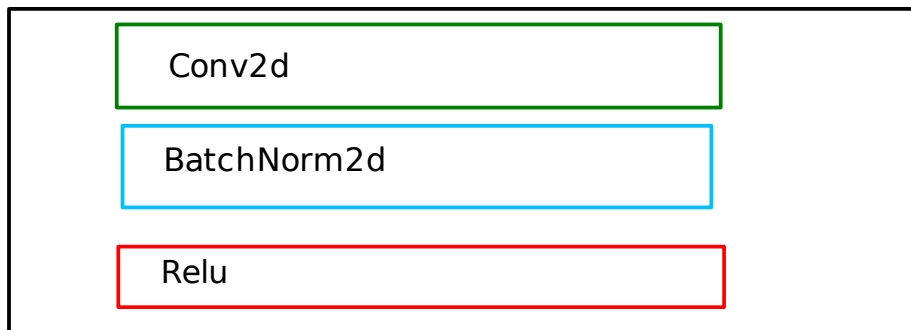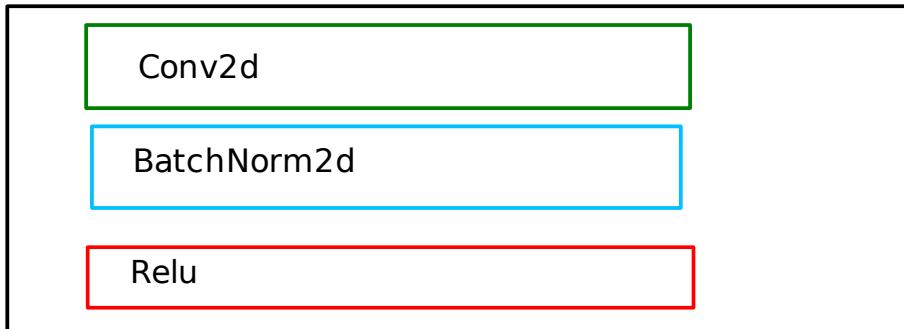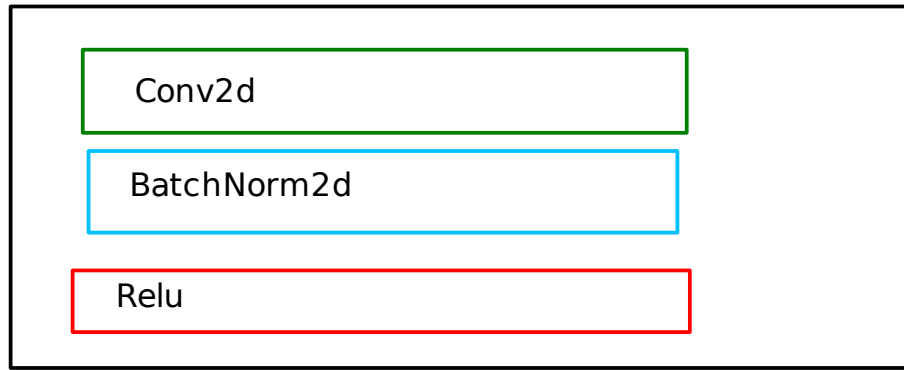Convst

input is batch_size * 1 channel * frames * input_features

(8,1,640,229)

Conv2d

BatchNorm2d

Relu

Conv2d

BatchNorm2d

Relu

MaxPool

Dropout

Conv2d

BatchNorm2d

Relu

MaxPool

Dropout

FC

Dropout

(8, 640, 768)