WakeupTriggerConvLSTM2s(nn.Module):

convulution + RELU

permute + reshape

LSTM Layer

batch normalization

fully Connected

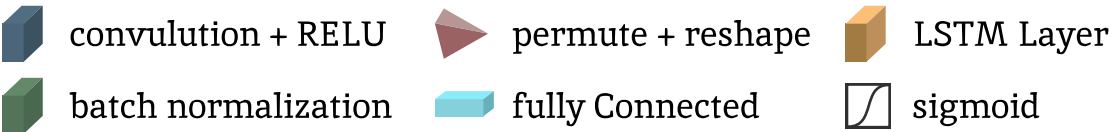sigmoid

Input Spectogram:
(Batch, Channel, Height, Width) -> 32 Batches while training, 1 Channel (mono), 128 Height x 256 Width

conv1

conv2

conv3

reshape

LSTM Block

fc1

fc2

sig1

[1, H, W]

[32, H/4, W/4]

[64, H/16, W/16]

[128, H/32, W/32]

[H/32, 128 * W/32]

[H/32, 8]

[32]

[1]

[1]