```
ConvolutionBackward0 AccumulateGrad AccumulateGrad
                                                  0.4.depthwise.weight (32, 1, 3, 3)
                 MaxPool2DWithIndicesBackward0 AccumulateGrad 0.4.pointwise.weight (64, 32, 1, 1)
                                                                       ConvolutionBackward0 AccumulateGrad AccumulateGrad
                                                                                             CudnnBatchNormBackward0 0.7.conv1.weight (64, 64, 3, 3)
                                                                                                                                                     0.7.bn1.weight (64) (64)
                                                                                                                                0.7.conv2.weight (64, 64, 3, 3)
                                                                                                    0.8.fc1.bias
(4)
                                                                                     AdaptiveAvgPool2DBackward0 AdaptiveAvgPool2DBackward0
                                                                                                                                          0.9.bn1.weight (8) (8)
                                                                                                                                 0.9.conv_h.weight (64, 8, 1, 1) (64, 8, 1, 1)
                                                                                                          SigmoidBackward0 SigmoidBackward0
                                                                                                                                                     0.10.depthwise.weight (64, 1, 3, 3)
                                                                                                                                                                           0.10.pointwise.weight (128, 64, 1, 1)
                                                                                                                                                      AccumulateGrad
                                                                                                                                                                        ConvolutionBackward0 AccumulateGrad AccumulateGrad
                                                                                                                                          0.13.conv_dil1.weight (128, 128, 3, 3)
                                                                                                                                                                     0.13.conv_dil1.bias (128)
                                                                                                                                                                                               CudnnBatchNormBackward0 0.13.conv_dil2.weight (128, 128, 3, 3)
                                                                                                                                                                                                                                                                                0.13.conv_dil3.weight (128, 128, 3, 3)
                                                                                                                                                                                                                                                      AccumulateGrad
                                                                                                                                                                                                                         ConvolutionBackward0

ConvolutionBackward0

ConvolutionBackward0

ConvolutionBackward0

ConvolutionBackward0

ConvolutionBackward0

ConvolutionBackward0

ConvolutionBackward0

ConvolutionBackward0
                                                                                                                                                                                                                                                                                  AccumulateGrad
                                                                                                                                                                                                                                                                                                      AccumulateGrad 0.14.conv1.weight (128, 128, 3, 3)
                                                                                                                                                                                                                                                                                                                                              0.14.bn1.weight (128) (128)
                                                                                                                                                                                                                                                                                                                       AccumulateGrad
                                                                                                                                                                                                                                                                                                                    ConvolutionBackward0 AccumulateGrad AccumulateGrad
                                                                                                                                                                                                                                                                                                                     0.14.conv2.weight (128, 128, 3, 3)
                                                                                                                                                                                                                                                                                                  0.14.bn2.weight (128)
                                                                                                                                                                                                                                                                                                                      AccumulateGrad NativeDropoutBackward0 0.14.bn2.bias (128)
                                                                                                                                                                                                                                                                                                                           0.16.depthwise.weight (128, 1, 3, 3)
                                                                                                                                                                                                                                                                                                                                        0.19.conv_dil1.bias | CudnnBatchNormBackward0 | 0.19.conv_dil2.weight (256, 256, 3, 3) | 0.19.conv_dil2.bias (256, 256, 3, 3)
                                                                                                                                                                                                                                                                                                                                                                                            ConvolutionBackward0 ConvolutionBackward0 (256, 768, 1, 1) 0.19.fuse.bias (256)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     AccumulateGrad

AccumulateGrad

0.20.conv1.weight (256, 256, 3, 3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0.20.bn1.weight (256)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ConvolutionBackward0 AccumulateGrad AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0.20.conv2.weight (256, 256, 3, 3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.21.fc1.weight (16, 256, 1, 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0.21.fc2.weight (256, 16, 1, 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NativeDropoutBackward0 AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0.22.depthwise.weight (256, 1, 3, 3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0.22.pointwise.weight (512, 256, 1, 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ConvolutionBackward0 AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ConvolutionBackward0 AccumulateGrad AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CudnnBatchNormBackward0 0.25.conv1.weight (512, 512, 3, 3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0.25.bn1.weight (512) (512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.25.conv2.weight (512, 512, 3, 3) ReluBackward0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0.26.channel_att.fc1.weight (32, 512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0.26.channel_att.fc1.bias (32)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0.26.channel_att.fc2.weight (512, 32)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MeanBackward1 TBackward0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0.26.channel_att.fc2.bias (512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.26.spatial_att.conv.weight (1, 2, 7, 7)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.0.transformer_encoder.layers.0.self_attn.in_proj_weight (1536, 512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1.0.transformer_encoder.layers.0.self_attn.in_proj_bias (1536)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SelectBackward0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.0.transformer_encoder.layers.0.self_attn.out_proj.weight (512, 512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1.0.transformer_encoder.layers.0.self_attn.out_proj.bias (512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.0.transformer_encoder.layers.0.norm1.weight (512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.0.transformer_encoder.layers.0.norm1.bias (512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1.0.transformer_encoder.layers.0.linear1.weight (2048, 512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1.0.transformer_encoder.layers.0.linear1.bias (2048)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.0.transformer_encoder.layers.0.linear2.weight (512, 2048)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.0.transformer_encoder.layers.0.linear2.bias (512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.0.transformer_encoder.layers.0.norm2.weight (512)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1.0.transformer_encoder.layers.0.norm2.bias
```