```
AddBackward0 AccumulateGrad
                                                                                                                                                                                                                 NativeLayerNormBackward0 stages.1.blocks.1.mlp.fc1.weight (320, 80, 1, 1) stages.1.blocks.1.mlp.fc1.bias (320)
                                                                                                                                                                                                                                ConvolutionBackward0 stages.1.blocks.1.mlp.fc2.weight (80, 320, 1, 1) stages.1.blocks.1.mlp.fc2.bias (80)
                                                                                                                                                                                                                                                                                                               NativeLayerNormBackward0 stages.2.blocks.0.mlp.fc1.weight (640, 160, 1, 1) stages.2.blocks.0.mlp.fc1.bias (640)
                                                                                                                                                                                                                                                                                                                                                                             ConvolutionBackward0 stages.2.blocks.1.norm.weight (160) stages.2.blocks.1.norm.bias
                                                                                                                                                                                                                                                                                                                                                                                                     ConvolutionBackward0 stages.2.blocks.1.mlp.fc2.weight (160, 640, 1, 1) stages.2.blocks.1.mlp.fc2.bias (160)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ConvolutionBackward0 stages.2.blocks.2.norm.weight (160) stages.2.blocks.2.norm.bias
                                                                                                                                                                                                                                                                                                                                                                                                                                                               NativeLayerNormBackward0 stages.2.blocks.2.mlp.fc1.weight (640, 160, 1, 1) stages.2.blocks.2.mlp.fc1.bias (640)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ConvolutionBackward0 stages.2.blocks.2.mlp.fc2.weight (160, 640, 1, 1) stages.2.blocks.2.mlp.fc2.bias (160)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ConvolutionBackward0 stages.2.blocks.3.norm.weight (160) stages.2.blocks.3.norm.bias (160)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NativeLayerNormBackward0 stages.2.blocks.3.mlp.fc1.weight (640, 160, 1, 1) stages.2.blocks.3.mlp.fc1.bias (640)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ConvolutionBackward0 stages.2.blocks.3.mlp.fc2.weight (160, 640, 1, 1) stages.2.blocks.3.mlp.fc2.bias (160)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ConvolutionBackward0 stages.2.blocks.4.norm.weight (160) stages.2.blocks.4.norm.bias
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NativeLayerNormBackward0 stages.2.blocks.4.mlp.fc1.weight (640, 160, 1, 1) stages.2.blocks.4.mlp.fc1.bias (640)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ConvolutionBackward0 stages.2.blocks.4.mlp.fc2.weight (160, 640, 1, 1) stages.2.blocks.4.mlp.fc2.bias (160)
```