```
bnl.bias
(64)
           AccumulateGrad
         ConvolutionBackward0
                                AccumulateGrad
                                                  AccumulateGrad
                                                       model_blocks.0.conv1.weight
                            NativeBatchNormBackward0
                                                             (64, 64, 3, 3)
                                                                         model_blocks.0.bn1.weight
                                                                                                     model_blocks.0.bn1.bias
                                      ReluBackward0
                                                      AccumulateGrad
                                                                            AccumulateGrad
                                                                                                    AccumulateGrad
                                                     ConvolutionBackward0
                                                   NativeBatchNormBackward0 model_blocks.0.conv2.weight (64, 64, 3, 3)
                                                                                            model_blocks.0.bn2.weight
                                                                                                                        model_blocks.0.bn2.bias
                                                        ReluBackward0
                                                                         AccumulateGrad
                                                                                                                                (64)
                                                                                                                      AccumulateGrad
                                                                        ConvolutionBackward0
                                                                                               AccumulateGrad
                                                                               NativeBatchNormBackward0
                                                                         model_blocks.1.conv1.weight (64, 64, 3, 3)
                                                         AddBackward0
                                                                                                                          model_blocks.1.bn1.bias
                                                                                               model_blocks.1.bn1.weight
                                                           ReluBackward0
                                                                            AccumulateGrad
                                                                                                                         AccumulateGrad
                                                                                                  AccumulateGrad
                                                                         ConvolutionBackward0
                                                                        NativeBatchNormBackward0
                                                                                                   model_blocks.1.conv2.weight (64, 64, 3, 3)
                                                                                                                                           model_blocks.1.bn2.bias
                                                                                                                 model_blocks.1.bn2.weight (64)
                                                                             ReluBackward0
                                                                                             AccumulateGrad
                                                                                                                                                     (64)
                                                                                                                    AccumulateGrad
                                                                                                                                           AccumulateGrad
                                                                                             ConvolutionBackward0
                                                                                                    NativeBatchNormBackward0
                                                 model_blocks.2.shortcut.0.weight (128, 64, 1, 1)
                                                                                                   model_blocks.2.conv1.weight (128, 64, 3, 3)
                                                                                   AddBackward0
                                 model_blocks.2.shortcut.1.bias (128)
                                                                                                                      model_blocks.2.bn1.weight (128)
                                                                                                                                                 model_blocks.2.bn1.bias
model_blocks.2.shortcut.1.weight
                                                                 AccumulateGrad ReluBackward0
                                                                                                    AccumulateGrad
         (128)
                                                                                                                                                         (128)
                                                              ConvolutionBackward0
                                                                                                 ConvolutionBackward0
                                                                                                                                                AccumulateGrad
               AccumulateGrad
                                           AccumulateGrad
                                                                                                                         AccumulateGrad
                                                                                                          NativeBatchNormBackward0 model_blocks.2.conv2.weight (128, 128, 3, 3)
                                                 NativeBatchNormBackward0
                                                                                                                                                  model_blocks.2.bn2.weight (128)
                                                                                                                                                                             model_blocks.2.bn2.bias
                                                                                                                               AccumulateGrad
                                                                                                               ReluBackward0
                                                                                                                                                                                      (128)
                                                                                                                              ConvolutionBackward0
                                                                                                                                                                             AccumulateGrad
                                                                                                                                                      AccumulateGrad
                                                                                                                                      NativeBatchNormBackward0
                                                                                                                       model_blocks.3.conv1.weight (128, 128, 3, 3)
                                                                                                       AddBackward0
                                                                                                                                                                       model_blocks.3.bn1.bias
                                                                                                                                           ReluBackward0
                                                                                                                          AccumulateGrad
                                                                                                                                                                                (128)
                                                                                                                       ConvolutionBackward0
                                                                                                                                                AccumulateGrad
                                                                                                                                                                       AccumulateGrad
                                                                                                                                                           model_blocks.3.conv2.weight (128, 128, 3, 3)
                                                                                                                               NativeBatchNormBackward0
                                                                                                                                                                       model_blocks.3.bn2.weight (128)
                                                                                                                                                                                                   model_blocks.3.bn2.bias
                                                                                                                                                    AccumulateGrad
                                                                                                                                                                                                           (128)
                                                                                                                                                   ConvolutionBackward0
                                                                                                                                                                           AccumulateGrad
                                                                                                                                                                                                  AccumulateGrad
                                                                                                                                                           NativeBatchNormBackward0
                                                                                                                                                      model_blocks.4.conv1.weight (256, 128, 3, 3)
                                                                                                    model_blocks.4.shortcut.0.weight
                                                                                                                                       AddBackward0
                                                                                                           (256, 128, 1, 1)
                                                model_blocks.4.shortcut.1.weight
                                                                                                                                                                        model_blocks.4.bn1.weight
                                                                                   model_blocks.4.shortcut.1.bias
                                                                                                                                                                                                    model_blocks.4.bn1.bias
                                                                                                                   AccumulateGrad ReluBackward0 AccumulateGrad
                                                                                                                                                                                                            (256)
                                                                                                                 ConvolutionBackward0
                                                                    AccumulateGrad
                                                                                             AccumulateGrad
                                                                                                                                                    ConvolutionBackward0
                                                                                                                                                                           AccumulateGrad
                                                                                                                                                                                                   AccumulateGrad
                                                                                                                                                            NativeBatchNormBackward0
                                                                                                                                                                                        model_blocks.4.conv2.weight
                                                                                                   NativeBatchNormBackward0
                                                                                                                                                                                            <sup>-</sup>(256, 256, 3, 3)
                                                                                                                                                                                                    model_blocks.4.bn2.weight
                                                                                                                                                                                                                                model_blocks.4.bn2.bias
                                                                                                                                                                  ReluBackward0
                                                                                                                                                                                  AccumulateGrad
                                                                                                                                                                                                                                        (256)
                                                                                                                                                                                                             (256)
                                                                                                                                                                                                                               AccumulateGrad
                                                                                                                                                                                ConvolutionBackward0
                                                                                                                                                                                                        AccumulateGrad
                                                                                                                                                                                        NativeBatchNormBackward0
                                                                                                                                                                         model_blocks.5.conv1.weight (256, 256, 3, 3)
                                                                                                                                                          AddBackward0
                                                                                                                                                          ReluBackward0
                                                                                                                                                                                              model_blocks.5.bn1.weight
                                                                                                                                                                                                                          model_blocks.5.bn1.bias
                                                                                                                                                                            AccumulateGrad
                                                                                                                                                                                                                                   (256)
                                                                                                                                                                          ConvolutionBackward0
                                                                                                                                                                                                  AccumulateGrad
                                                                                                                                                                                                                         AccumulateGrad
                                                                                                                                                                                                    model_blocks.5.conv2.weight
                                                                                                                                                                         NativeBatchNormBackward0
                                                                                                                                                                                                         (256, 256, 3, 3)
                                                                                                                                                                                                                 model_blocks.5.bn2.weight
                                                                                                                                                                                                                                             model_blocks.5.bn2.bias
                                                                                                                                                                              ReluBackward0
                                                                                                                                                                                               AccumulateGrad
                                                                                                                                                                                                                                                     (256)
                                                                                                                                                                                           ConvolutionBackward0 AccumulateGrad
                                                                                                                                                                                                                                          AccumulateGrad
                                                                                                                                                                                                     NativeBatchNormBackward0
                                                                                                                                                  model_blocks.6.shortcut.0.weight (128, 256, 1, 1)
                                                                                                                                                                                                    model_blocks.6.conv1.weight (128, 256, 3, 3)
                                                                                                                                                                                  AddBackward0
                                                                                                                                 model_blocks.6.shortcut.1.bias (128)
                                                                                                                                                                                                                      model_blocks.6.bn1.weight
                                                                                               model_blocks.6.shortcut.1.weight
                                                                                                                                                                 AccumulateGrad ReluBackward0 AccumulateGrad
                                                                                                                                                                                                                                                  model_blocks.6.bn1.bias
                                                                                                          (128)
                                                                                                                                                                                                                                                       (128)
                                                                                                                                                              ConvolutionBackward0
                                                                                                                  AccumulateGrad
                                                                                                                                           AccumulateGrad
                                                                                                                                                                                                  ConvolutionBackward0
                                                                                                                                                                                                                          AccumulateGrad
                                                                                                                                                                                                                                                 AccumulateGrad
                                                                                                                                                                                                                                     model_blocks.6.conv2.weight (128, 128, 3, 3)
                                                                                                                                                                                                          NativeBatchNormBackward0
                                                                                                                                                NativeBatchNormBackward0
                                                                                                                                                                                                                                                  model_blocks.6.bn2.weight
                                                                                                                                                                                                                                                                             model_blocks.6.bn2.bias
                                                                                                                                                                                                               ReluBackward0
                                                                                                                                                                                                                                AccumulateGrad
                                                                                                                                                                                                                                                                                      (128)
                                                                                                                                                                                                                              ConvolutionBackward0
                                                                                                                                                                                                                                                     AccumulateGrad
                                                                                                                                                                                                                                                                            AccumulateGrad
                                                                                                                                                                                                                                      NativeBatchNormBackward0
                                                                                                                                                                                                                       model_blocks.7.conv1.weight (128, 128, 3, 3)
                                                                                                                                                                                                        AddBackward0
                                                                                                                                                                                                                                            model_blocks.7.bn1.weight
                                                                                                                                                                                                                                                                        model_blocks.7.bn1.bias
                                                                                                                                                                                                         ReluBackward0
                                                                                                                                                                                                                          AccumulateGrad
                                                                                                                                                                                                                                                      (128)
                                                                                                                                                                                                                                                                              (128)
                                                                                                                                                                                                                                                                      AccumulateGrad
                                                                                                                                                                                                                        ConvolutionBackward0
                                                                                                                                                                                                                                                AccumulateGrad
                                                                                                                                                                                                                                                          model_blocks.7.conv2.weight (128, 128, 3, 3)
                                                                                                                                                                                                                               NativeBatchNormBackward0
                                                                                                                                                                                                                                                                       model_blocks.7.bn2.weight
                                                                                                                                                                                                                                                                                                   model_blocks.7.bn2.bias
                                                                                                                                                                                                                                                     AccumulateGrad
                                                                                                                                                                                                                                     ReluBackward0
                                                                                                                                                                                                                                                                                                            (128)
                                                                                                                                                                                                                                                   ConvolutionBackward0
                                                                                                                                                                                                                                                                                                 AccumulateGrad
                                                                                                                                                                                                                                                                          AccumulateGrad
                                                                                                                                                                                                                                                           NativeBatchNormBackward0
                                                                                                                                                                                                                                      AddBackward0
                                                                                                                                                                                                                                                     model_blocks.8.conv1.weight
                                                                                                                                                                                                    model_blocks.8.shortcut.0.weight
                                                                                                                                                                                                           (64, 128, 1, 1)
                                                                                                                                                                                                                                                          (64, 128, 3, 3)
                                                                                                                                                                                  model_blocks.8.bn1.weight
                                                                                                                                                  model_blocks.8.shortcut.1.weight
                                                                                                                                                                                                                                                                                                    model_blocks.8.bn1.bias
                                                                                                                                                              (64)
                                                                                                                                                                                                                                                    ConvolutionBackward0
                                                                                                                                                                 AccumulateGrad
                                                                                                                                                                                             AccumulateGrad
                                                                                                                                                                                                                ConvolutionBackward0
                                                                                                                                                                                                                                                                           AccumulateGrad
                                                                                                                                                                                                                                                                                                   AccumulateGrad
                                                                                                                                                                                                                                                            NativeBatchNormBackward0 model_blocks.8.conv2.weight (64, 64, 3, 3)
                                                                                                                                                                                                    NativeBatchNormBackward0
                                                                                                                                                                                                                                                                                                     model_blocks.8.bn2.weight
                                                                                                                                                                                                                                                                                                                                model_blocks.8.bn2.bias
                                                                                                                                                                                                                                                                 ReluBackward0 AccumulateGrad
                                                                                                                                                                                                                                                                                                                                        (64)
                                                                                                                                                                                                                                                                                ConvolutionBackward0
                                                                                                                                                                                                                                                                                                        AccumulateGrad
                                                                                                                                                                                                                                                                                                                               AccumulateGrad
                                                                                                                                                                                                                                                                                        NativeBatchNormBackward0
                                                                                                                                                                                                                                                                          model_blocks.9.conv1.weight (64, 64, 3, 3)
                                                                                                                                                                                                                                                         AddBackward0
                                                                                                                                                                                                                                                                                              model_blocks.9.bn1.weight
                                                                                                                                                                                                                                                                                                                          model_blocks.9.bn1.bias
                                                                                                                                                                                                                                                           ReluBackward0
                                                                                                                                                                                                                                                                            AccumulateGrad
                                                                                                                                                                                                                                                                                                  AccumulateGrad
                                                                                                                                                                                                                                                                                                                         AccumulateGrad
                                                                                                                                                                                                                                                                          ConvolutionBackward0
                                                                                                                                                                                                                                                                                                            model_blocks.9.conv2.weight (64, 64, 3, 3)
                                                                                                                                                                                                                                                                                  NativeBatchNormBackward0
                                                                                                                                                                                                                                                                                                                          model_blocks.9.bn2.weight
                                                                                                                                                                                                                                                                                                                                                     model_blocks.9.bn2.bias
                                                                                                                                                                                                                                                                                      ReluBackward0
                                                                                                                                                                                                                                                                                                        AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                              (64)
                                                                                                                                                                                                                                                                                                      ConvolutionBackward0
                                                                                                                                                                                                                                                                                                                                                    AccumulateGrad
                                                                                                                                                                                                                                                                                                                              AccumulateGrad
                                                                                                                                                                                                                                                                                                             NativeBatchNormBackward0
                                                                                                                                                                                                                                                     model_blocks.10.shortcut.0.weight
                                                                                                                                                                                                                                                                                                        model_blocks.10.conv1.weight
                                                                                                                                                                                                                                                                                        AddBackward0
                                                                                                                                                                                                                                                             (128, 64, 1, 1)
                                                                                                                                                                                                                                                                                                              (128, 64, 3, 3)
                                                                                                                                                                                                   model_blocks.10.shortcut.1.weight
                                                                                                                                                                                                                                     model_blocks.10.shortcut.1.bias
                                                                                                                                                                                                                                                                                                                           model_blocks.10.bn1.weight
                                                                                                                                                                                                                                                                                                                                                       model_blocks.10.bn1.bias
                                                                                                                                                                                                                                                                                      ReluBackward0
                                                                                                                                                                                                                                                                                                        AccumulateGrad
                                                                                                                                                                                                                                                                      AccumulateGrad
                                                                                                                                                                                                             (128)
                                                                                                                                                                                                                                               (128)
                                                                                                                                                                                                                                                                                                                                     (128)
                                                                                                                                                                                                                                                                                                                                                                (128)
                                                                                                                                                                                                                                                                   ConvolutionBackward0
                                                                                                                                                                                                                     AccumulateGrad
                                                                                                                                                                                                                                               AccumulateGrad
                                                                                                                                                                                                                                                                                                      ConvolutionBackward0
                                                                                                                                                                                                                                                                                                                              AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                      AccumulateGrad
                                                                                                                                                                                                                                                    NativeBatchNormBackward0
                                                                                                                                                                                                                                                                                                               NativeBatchNormBackward0
                                                                                                                                                                                                                                                                                                                                          model_blocks.10.conv2.weight (128, 128, 3, 3)
                                                                                                                                                                                                                                                                                                                                                       model_blocks.10.bn2.weight
                                                                                                                                                                                                                                                                                                                                                                                    model_blocks.10.bn2.bias
                                                                                                                                                                                                                                                                                                                    ReluBackward0
                                                                                                                                                                                                                                                                                                                                     AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                                            (128)
                                                                                                                                                                                                                                                                                                                                                                                   AccumulateGrad
                                                                                                                                                                                                                                                                                                                                   ConvolutionBackward0
                                                                                                                                                                                                                                                                                                                                                           AccumulateGrad
                                                                                                                                                                                                                                                                                                                                          NativeBatchNormBackward0
                                                                                                                                                                                                                                                                                    model_blocks.11.shortcut.0.weight
                                                                                                                                                                                                                                                                                                                                       model_blocks.11.conv1.weight
                                                                                                                                                                                                                                                                                                                        AddBackward0
                                                                                                                                                                                                                                                                                            (256, 128, 1, 1)
                                                                                                                                                                                                                                                                                                                                            (256, 128, 3, 3)
                                                                                                                                                                                                                                 model_blocks.11.shortcut.1.weight
                                                                                                                                                                                                                                                                   model_blocks.11.shortcut.1.bias
                                                                                                                                                                                                                                                                                                                                                         model_blocks.11.bn1.weight
                                                                                                                                                                                                                                                                                                                                                                                      model_blocks.11.bn1.bias
                                                                                                                                                                                                                                                                                                     AccumulateGrad
                                                                                                                                                                                                                                                                                                                                      AccumulateGrad
                                                                                                                                                                                                                                                                                                                     ReluBackward0
                                                                                                                                                                                                                                              (256)
                                                                                                                                                                                                                                                                                (256)
                                                                                                                                                                                                                                                                                                                                                                    (256)
                                                                                                                                                                                                                                                                                                                                                                                               (256)
                                                                                                                                                                                                                                                                                                                                                                                    AccumulateGrad
                                                                                                                                                                                                                                                                             AccumulateGrad
                                                                                                                                                                                                                                                    AccumulateGrad
                                                                                                                                                                                                                                                                                                  ConvolutionBackward0
                                                                                                                                                                                                                                                                                                                                    ConvolutionBackward0
                                                                                                                                                                                                                                                                                                                                                             AccumulateGrad
                                                                                                                                                                                                                                                                                   NativeBatchNormBackward0
                                                                                                                                                                                                                                                                                                                                                                         model_blocks.11.conv2.weight (256, 256, 3, 3)
                                                                                                                                                                                                                                                                                                                                              NativeBatchNormBackward0
                                                                                                                                                                                                                                                                                                                                                                                     model_blocks.11.bn2.weight
                                                                                                                                                                                                                                                                                                                                                                                                                  model_blocks.11.bn2.bias
                                                                                                                                                                                                                                                                                                                                                   ReluBackward0
                                                                                                                                                                                                                                                                                                                                                                   AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                 ConvolutionBackward0
                                                                                                                                                                                                                                                                                                                                                                                                                 AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                                          AccumulateGrad
                                                                                                                                                                                                                                                                                                                                                                         NativeBatchNormBackward0
                                                                                                                                                                                                                                                                                                                                           AddBackward0
                                                                                                                                                                                                                                                                                                                                                            linear.weight
(10, 256)
                                                                                                                                                                                                                                                                                                                                         ReluBackward0 AccumulateGrad
                                                                                                                                                                                                                                                                                                                        AccumulateGrad ViewBackward0
```

conv1.weight (64, 3, 3)