APPENDIX

Ⅰ Complexity information for ResNet

ResNet18:

ResNet(

11.69 M, 100.000% Params, 2.38 GMac, 100.000% MACs,

(conv1): Conv2d(0.009 M, 0.080% Params, 0.154 GMac, 6.477% MACs, 3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)

(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.002 GMac, 0.088% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.044% MACs, inplace=True)

(maxpool): MaxPool2d(0.0 M, 0.000% Params, 0.001 GMac, 0.044% MACs, kernel_size=3, stride=2, padding=1, dilation=1, ceil_mode=False)

(layer1): Sequential(

0.148 M, 1.266% Params, 0.607 GMac, 25.511% MACs,

(0): BasicBlock(

0.074 M, 0.633% Params, 0.304 GMac, 12.756% MACs,

(conv1): Conv2d(0.037 M, 0.315% Params, 0.151 GMac, 6.345% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)

(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.022% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.022% MACs, inplace=True)

(conv2): Conv2d(0.037 M, 0.315% Params, 0.151 GMac, 6.345% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)

(bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.022% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

)

(1): BasicBlock(

0.074 M, 0.633% Params, 0.304 GMac, 12.756% MACs,

(conv1): Conv2d(0.037 M, 0.315% Params, 0.151 GMac, 6.345% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)

(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.022% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.022% MACs, inplace=True)

(conv2): Conv2d(0.037 M, 0.315% Params, 0.151 GMac, 6.345% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)

(bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.022% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

)

)

(layer2): Sequential(

0.526 M, 4.496% Params, 0.539 GMac, 22.636% MACs,

(0): BasicBlock(

0.23 M, 1.969% Params, 0.236 GMac, 9.914% MACs,

(conv1): Conv2d(0.074 M, 0.631% Params, 0.075 GMac, 3.172% MACs, 64, 128,

kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.002% Params, 0.0 GMac, 0.011% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.011% MACs, inplace=True)
(conv2): Conv2d(0.147 M, 1.261% Params, 0.151 GMac, 6.345% MACs, 128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.0 M, 0.002% Params, 0.0 GMac, 0.011% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(downsample): Sequential(
0.008 M, 0.072% Params, 0.009 GMac, 0.364% MACs,
(0): Conv2d(0.008 M, 0.070% Params, 0.008 GMac, 0.352% MACs, 64, 128, kernel_size=(1, 1), stride=(2, 2), bias=False)
(1): BatchNorm2d(0.0 M, 0.002% Params, 0.0 GMac, 0.011% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(1): BasicBlock(
0.295 M, 2.527% Params, 0.303 GMac, 12.723% MACs,
(conv1): Conv2d(0.147 M, 1.261% Params, 0.151 GMac, 6.345% MACs, 128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.002% Params, 0.0 GMac, 0.011% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.011% MACs, inplace=True)
(conv2): Conv2d(0.147 M, 1.261% Params, 0.151 GMac, 6.345% MACs, 128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.0 M, 0.002% Params, 0.0 GMac, 0.011% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(layer3): Sequential(
2.1 M, 17.962% Params, 0.538 GMac, 22.598% MACs,
(0): BasicBlock(
0.919 M, 7.862% Params, 0.235 GMac, 9.892% MACs,
(conv1): Conv2d(0.295 M, 2.523% Params, 0.075 GMac, 3.172% MACs, 128, 256, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.006% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.006% MACs, inplace=True)
(conv2): Conv2d(0.59 M, 5.046% Params, 0.151 GMac, 6.345% MACs, 256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.006% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(downsample): Sequential(
0.033 M, 0.285% Params, 0.009 GMac, 0.358% MACs,

(0): Conv2d(0.033 M, 0.280% Params, 0.008 GMac, 0.352% MACs, 128, 256, kernel_size=(1, 1), stride=(2, 2), bias=False)
(1): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.006% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(1): BasicBlock(
1.181 M, 10.100% Params, 0.302 GMac, 12.706% MACs,
(conv1): Conv2d(0.59 M, 5.046% Params, 0.151 GMac, 6.345% MACs, 256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.006% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.006% MACs, inplace=True)
(conv2): Conv2d(0.59 M, 5.046% Params, 0.151 GMac, 6.345% MACs, 256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.006% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(layer4): Sequential(
8.394 M, 71.806% Params, 0.537 GMac, 22.578% MACs,
(0): BasicBlock(
3.673 M, 31.422% Params, 0.235 GMac, 9.881% MACs,
(conv1): Conv2d(1.18 M, 10.092% Params, 0.075 GMac, 3.172% MACs, 256, 512, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.009% Params, 0.0 GMac, 0.003% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.003% MACs, inplace=True)
(conv2): Conv2d(2.359 M, 20.183% Params, 0.151 GMac, 6.345% MACs, 512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.009% Params, 0.0 GMac, 0.003% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(downsample): Sequential(
0.132 M, 1.130% Params, 0.008 GMac, 0.355% MACs,
(0): Conv2d(0.131 M, 1.121% Params, 0.008 GMac, 0.352% MACs, 256, 512, kernel_size=(1, 1), stride=(2, 2), bias=False)
(1): BatchNorm2d(0.001 M, 0.009% Params, 0.0 GMac, 0.003% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(1): BasicBlock(
4.721 M, 40.384% Params, 0.302 GMac, 12.698% MACs,
(conv1): Conv2d(2.359 M, 20.183% Params, 0.151 GMac, 6.345% MACs, 512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)

(bn1): BatchNorm2d(0.001 M, 0.009% Params, 0.0 GMac, 0.003% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.003% MACs, inplace=True)
(conv2): Conv2d(2.359 M, 20.183% Params, 0.151 GMac, 6.345% MACs, 512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.009% Params, 0.0 GMac, 0.003% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(avgpool): AdaptiveAvgPool2d(0.0 M, 0.000% Params, 0.0 GMac, 0.001% MACs, output_size=(1, 1))
(fc): Linear(0.513 M, 4.389% Params, 0.001 GMac, 0.022% MACs, in_features=512, out_features=1000, bias=True)
)
Computational complexity: 2379830248.0
Number of parameters: 11689512

ResNet34:

ResNet(
21.798 M, 100.000% Params, 4.801 GMac, 100.000% MACs,
(conv1): Conv2d(0.009 M, 0.043% Params, 0.154 GMac, 3.211% MACs, 3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.002 GMac, 0.044% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.022% MACs, inplace=True)
(maxpool): MaxPool2d(0.0 M, 0.000% Params, 0.001 GMac, 0.022% MACs, kernel_size=3, stride=2, padding=1, dilation=1, ceil_mode=False)
(layer1): Sequential(
0.222 M, 1.018% Params, 0.911 GMac, 18.970% MACs,
(0): BasicBlock(
0.074 M, 0.339% Params, 0.304 GMac, 6.323% MACs,
(conv1): Conv2d(0.037 M, 0.169% Params, 0.151 GMac, 3.145% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.011% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.011% MACs, inplace=True)
(conv2): Conv2d(0.037 M, 0.169% Params, 0.151 GMac, 3.145% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.011% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
(1): BasicBlock(
0.074 M, 0.339% Params, 0.304 GMac, 6.323% MACs,
(conv1): Conv2d(0.037 M, 0.169% Params, 0.151 GMac, 3.145% MACs, 64, 64,

kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.011% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.011% MACs, inplace=True)
(conv2): Conv2d(0.037 M, 0.169% Params, 0.151 GMac, 3.145% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.011% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
(2): BasicBlock(
0.074 M, 0.339% Params, 0.304 GMac, 6.323% MACs,
(conv1): Conv2d(0.037 M, 0.169% Params, 0.151 GMac, 3.145% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.011% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.011% MACs, inplace=True)
(conv2): Conv2d(0.037 M, 0.169% Params, 0.151 GMac, 3.145% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.011% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(layer2): Sequential(
1.116 M, 5.122% Params, 1.144 GMac, 23.835% MACs,
(0): BasicBlock(
0.23 M, 1.056% Params, 0.236 GMac, 4.915% MACs,
(conv1): Conv2d(0.074 M, 0.338% Params, 0.075 GMac, 1.573% MACs, 64, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.005% MACs, inplace=True)
(conv2): Conv2d(0.147 M, 0.676% Params, 0.151 GMac, 3.145% MACs, 128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(downsample): Sequential(
0.008 M, 0.039% Params, 0.009 GMac, 0.180% MACs,
(0): Conv2d(0.008 M, 0.038% Params, 0.008 GMac, 0.175% MACs, 64, 128, kernel_size=(1, 1), stride=(2, 2), bias=False)
(1): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(1): BasicBlock(

```
        0.295 M, 1.355% Params, 0.303 GMac, 6.307% MACs,
        (conv1): Conv2d(0.147 M, 0.676% Params, 0.151 GMac, 3.145% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.005% MACs, inplace=True)
        (conv2): Conv2d(0.147 M, 0.676% Params, 0.151 GMac, 3.145% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
      )
      (2): BasicBlock(
        0.295 M, 1.355% Params, 0.303 GMac, 6.307% MACs,
        (conv1): Conv2d(0.147 M, 0.676% Params, 0.151 GMac, 3.145% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.005% MACs, inplace=True)
        (conv2): Conv2d(0.147 M, 0.676% Params, 0.151 GMac, 3.145% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
      )
      (3): BasicBlock(
        0.295 M, 1.355% Params, 0.303 GMac, 6.307% MACs,
        (conv1): Conv2d(0.147 M, 0.676% Params, 0.151 GMac, 3.145% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.005% MACs, inplace=True)
        (conv2): Conv2d(0.147 M, 0.676% Params, 0.151 GMac, 3.145% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (layer3): Sequential(
      6.822 M, 31.299% Params, 1.747 GMac, 36.397% MACs,
      (0): BasicBlock(
        0.919 M, 4.216% Params, 0.235 GMac, 4.904% MACs,
        (conv1): Conv2d(0.295 M, 1.353% Params, 0.075 GMac, 1.573% MACs, 128, 256,
kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
```

```
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.003% MACs, inplace=True)
        (conv2): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (downsample): Sequential(
          0.033 M, 0.153% Params, 0.009 GMac, 0.177% MACs,
          (0): Conv2d(0.033 M, 0.150% Params, 0.008 GMac, 0.175% MACs, 128, 256,
kernel_size=(1, 1), stride=(2, 2), bias=False)
          (1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        )
      )
      (1): BasicBlock(
        1.181 M, 5.417% Params, 0.302 GMac, 6.299% MACs,
        (conv1): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.003% MACs, inplace=True)
        (conv2): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
      )
      (2): BasicBlock(
        1.181 M, 5.417% Params, 0.302 GMac, 6.299% MACs,
        (conv1): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.003% MACs, inplace=True)
        (conv2): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
      )
      (3): BasicBlock(
        1.181 M, 5.417% Params, 0.302 GMac, 6.299% MACs,
        (conv1): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.003% MACs, inplace=True)
```

```
        (conv2): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
      )
      (4): BasicBlock(
        1.181 M, 5.417% Params, 0.302 GMac, 6.299% MACs,
        (conv1): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.003% MACs, inplace=True)
        (conv2): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
      )
      (5): BasicBlock(
        1.181 M, 5.417% Params, 0.302 GMac, 6.299% MACs,
        (conv1): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.003% MACs, inplace=True)
        (conv2): Conv2d(0.59 M, 2.706% Params, 0.151 GMac, 3.145% MACs, 256, 256,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.003% MACs, 256, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (layer4): Sequential(
      13.114 M, 60.164% Params, 0.84 GMac, 17.487% MACs,
      (0): BasicBlock(
        3.673 M, 16.851% Params, 0.235 GMac, 4.898% MACs,
        (conv1): Conv2d(1.18 M, 5.412% Params, 0.075 GMac, 1.573% MACs, 256, 512,
kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.001 M, 0.005% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.001% MACs, inplace=True)
        (conv2): Conv2d(2.359 M, 10.824% Params, 0.151 GMac, 3.145% MACs, 512, 512,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.001 M, 0.005% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (downsample): Sequential(
```

0.132 M, 0.606% Params, 0.008 GMac, 0.176% MACs,
(0): Conv2d(0.131 M, 0.601% Params, 0.008 GMac, 0.175% MACs, 256, 512, kernel_size=(1, 1), stride=(2, 2), bias=False)
(1): BatchNorm2d(0.001 M, 0.005% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(1): BasicBlock(
4.721 M, 21.657% Params, 0.302 GMac, 6.295% MACs,
(conv1): Conv2d(2.359 M, 10.824% Params, 0.151 GMac, 3.145% MACs, 512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.005% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.001% MACs, inplace=True)
(conv2): Conv2d(2.359 M, 10.824% Params, 0.151 GMac, 3.145% MACs, 512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.005% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
(2): BasicBlock(
4.721 M, 21.657% Params, 0.302 GMac, 6.295% MACs,
(conv1): Conv2d(2.359 M, 10.824% Params, 0.151 GMac, 3.145% MACs, 512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.005% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.001% MACs, inplace=True)
(conv2): Conv2d(2.359 M, 10.824% Params, 0.151 GMac, 3.145% MACs, 512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.005% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(avgpool): AdaptiveAvgPool2d(0.0 M, 0.000% Params, 0.0 GMac, 0.001% MACs, output_size=(1, 1))
(fc): Linear(0.513 M, 2.353% Params, 0.001 GMac, 0.011% MACs, in_features=512, out_features=1000, bias=True)
)
Computational complexity: 4800664552.0
Number of parameters: 21797672

ResNet50:

ResNet(
25.557 M, 100.000% Params, 5.383 GMac, 100.000% MACs,

(conv1): Conv2d(0.009 M, 0.037% Params, 0.154 GMac, 2.863% MACs, 3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.002 GMac, 0.039% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.019% MACs, inplace=True)
(maxpool): MaxPool2d(0.0 M, 0.000% Params, 0.001 GMac, 0.019% MACs, kernel_size=3, stride=2, padding=1, dilation=1, ceil_mode=False)
(layer1): Sequential(
0.216 M, 0.844% Params, 0.889 GMac, 16.508% MACs,
(0): Bottleneck(
0.075 M, 0.293% Params, 0.309 GMac, 5.737% MACs,
(conv1): Conv2d(0.004 M, 0.016% Params, 0.017 GMac, 0.312% MACs, 64, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.010% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(0.037 M, 0.144% Params, 0.151 GMac, 2.805% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.010% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(0.016 M, 0.064% Params, 0.067 GMac, 1.247% MACs, 64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(0.001 M, 0.002% Params, 0.002 GMac, 0.039% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.002 GMac, 0.029% MACs, inplace=True)
(downsample): Sequential(
0.017 M, 0.066% Params, 0.069 GMac, 1.286% MACs,
(0): Conv2d(0.016 M, 0.064% Params, 0.067 GMac, 1.247% MACs, 64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(1): BatchNorm2d(0.001 M, 0.002% Params, 0.002 GMac, 0.039% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(1): Bottleneck(
0.07 M, 0.275% Params, 0.29 GMac, 5.386% MACs,
(conv1): Conv2d(0.016 M, 0.064% Params, 0.067 GMac, 1.247% MACs, 256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.010% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(0.037 M, 0.144% Params, 0.151 GMac, 2.805% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.010% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(0.016 M, 0.064% Params, 0.067 GMac, 1.247% MACs, 64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)

```
        (bn3): BatchNorm2d(0.001 M, 0.002% Params, 0.002 GMac, 0.039% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.002 GMac, 0.029% MACs, inplace=True)
      )
      (2): Bottleneck(
        0.07 M, 0.275% Params, 0.29 GMac, 5.386% MACs,
        (conv1): Conv2d(0.016 M, 0.064% Params, 0.067 GMac, 1.247% MACs, 256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.010% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (conv2): Conv2d(0.037 M, 0.144% Params, 0.151 GMac, 2.805% MACs, 64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.010% MACs, 64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (conv3): Conv2d(0.016 M, 0.064% Params, 0.067 GMac, 1.247% MACs, 64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn3): BatchNorm2d(0.001 M, 0.002% Params, 0.002 GMac, 0.039% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.002 GMac, 0.029% MACs, inplace=True)
      )
    )
    (layer2): Sequential(
      1.22 M, 4.772% Params, 1.354 GMac, 25.150% MACs,
      (0): Bottleneck(
        0.379 M, 1.484% Params, 0.491 GMac, 9.123% MACs,
        (conv1): Conv2d(0.033 M, 0.128% Params, 0.134 GMac, 2.493% MACs, 256, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.001 GMac, 0.019% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (conv2): Conv2d(0.147 M, 0.577% Params, 0.151 GMac, 2.805% MACs, 128, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (conv3): Conv2d(0.066 M, 0.256% Params, 0.067 GMac, 1.247% MACs, 128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn3): BatchNorm2d(0.001 M, 0.004% Params, 0.001 GMac, 0.019% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.022% MACs, inplace=True)
        (downsample): Sequential(
          0.132 M, 0.517% Params, 0.135 GMac, 2.513% MACs,
          (0): Conv2d(0.131 M, 0.513% Params, 0.134 GMac, 2.493% MACs, 256, 512, kernel_size=(1, 1), stride=(2, 2), bias=False)
          (1): BatchNorm2d(0.001 M, 0.004% Params, 0.001 GMac, 0.019% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
```

```
        )
      )
      (1): Bottleneck(
        0.28 M, 1.096% Params, 0.288 GMac, 5.342% MACs,
        (conv1): Conv2d(0.066 M, 0.256% Params, 0.067 GMac, 1.247% MACs, 512, 128,
kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
        (conv2): Conv2d(0.147 M, 0.577% Params, 0.151 GMac, 2.805% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
        (conv3): Conv2d(0.066 M, 0.256% Params, 0.067 GMac, 1.247% MACs, 128, 512,
kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn3): BatchNorm2d(0.001 M, 0.004% Params, 0.001 GMac, 0.019% MACs, 512, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.015% MACs, inplace=True)
      )
      (2): Bottleneck(
        0.28 M, 1.096% Params, 0.288 GMac, 5.342% MACs,
        (conv1): Conv2d(0.066 M, 0.256% Params, 0.067 GMac, 1.247% MACs, 512, 128,
kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
        (conv2): Conv2d(0.147 M, 0.577% Params, 0.151 GMac, 2.805% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
        (conv3): Conv2d(0.066 M, 0.256% Params, 0.067 GMac, 1.247% MACs, 128, 512,
kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn3): BatchNorm2d(0.001 M, 0.004% Params, 0.001 GMac, 0.019% MACs, 512, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.015% MACs, inplace=True)
      )
      (3): Bottleneck(
        0.28 M, 1.096% Params, 0.288 GMac, 5.342% MACs,
        (conv1): Conv2d(0.066 M, 0.256% Params, 0.067 GMac, 1.247% MACs, 512, 128,
kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn1): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
        (conv2): Conv2d(0.147 M, 0.577% Params, 0.151 GMac, 2.805% MACs, 128, 128,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(0.0 M, 0.001% Params, 0.0 GMac, 0.005% MACs, 128, eps=1e-05,
momentum=0.1, affine=True, track_running_stats=True)
```

(conv3): Conv2d(0.066 M, 0.256% Params, 0.067 GMac, 1.247% MACs, 128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(0.001 M, 0.004% Params, 0.001 GMac, 0.019% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.015% MACs, inplace=True)
)
)
(layer3): Sequential(
7.098 M, 27.775% Params, 1.921 GMac, 35.682% MACs,
(0): Bottleneck(
1.512 M, 5.918% Params, 0.489 GMac, 9.081% MACs,
(conv1): Conv2d(0.131 M, 0.513% Params, 0.134 GMac, 2.493% MACs, 512, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.001 GMac, 0.010% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(0.59 M, 2.308% Params, 0.151 GMac, 2.805% MACs, 256, 256, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(0.002 M, 0.008% Params, 0.001 GMac, 0.010% MACs, 1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.001 GMac, 0.011% MACs, inplace=True)
(downsample): Sequential(
0.526 M, 2.059% Params, 0.135 GMac, 2.503% MACs,
(0): Conv2d(0.524 M, 2.051% Params, 0.134 GMac, 2.493% MACs, 512, 1024, kernel_size=(1, 1), stride=(2, 2), bias=False)
(1): BatchNorm2d(0.002 M, 0.008% Params, 0.001 GMac, 0.010% MACs, 1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(1): Bottleneck(
1.117 M, 4.371% Params, 0.286 GMac, 5.320% MACs,
(conv1): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(0.59 M, 2.308% Params, 0.151 GMac, 2.805% MACs, 256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)

(bn3): BatchNorm2d(0.002 M, 0.008% Params, 0.001 GMac, 0.010% MACs, 1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.007% MACs, inplace=True)
)
(2): Bottleneck(
1.117 M, 4.371% Params, 0.286 GMac, 5.320% MACs,
(conv1): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(0.59 M, 2.308% Params, 0.151 GMac, 2.805% MACs, 256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(0.002 M, 0.008% Params, 0.001 GMac, 0.010% MACs, 1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.007% MACs, inplace=True)
)
(3): Bottleneck(
1.117 M, 4.371% Params, 0.286 GMac, 5.320% MACs,
(conv1): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(0.59 M, 2.308% Params, 0.151 GMac, 2.805% MACs, 256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(0.002 M, 0.008% Params, 0.001 GMac, 0.010% MACs, 1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.007% MACs, inplace=True)
)
(4): Bottleneck(
1.117 M, 4.371% Params, 0.286 GMac, 5.320% MACs,
(conv1): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(0.59 M, 2.308% Params, 0.151 GMac, 2.805% MACs, 256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)

(bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(0.002 M, 0.008% Params, 0.001 GMac, 0.010% MACs, 1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.007% MACs, inplace=True)
)
(5): Bottleneck(
1.117 M, 4.371% Params, 0.286 GMac, 5.320% MACs,
(conv1): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(0.59 M, 2.308% Params, 0.151 GMac, 2.805% MACs, 256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.002% Params, 0.0 GMac, 0.002% MACs, 256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(0.262 M, 1.026% Params, 0.067 GMac, 1.247% MACs, 256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(0.002 M, 0.008% Params, 0.001 GMac, 0.010% MACs, 1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.007% MACs, inplace=True)
)
)
(layer4): Sequential(
14.965 M, 58.554% Params, 1.059 GMac, 19.678% MACs,
(0): Bottleneck(
6.04 M, 23.632% Params, 0.488 GMac, 9.060% MACs,
(conv1): Conv2d(0.524 M, 2.051% Params, 0.134 GMac, 2.493% MACs, 1024, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.005% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(2.359 M, 9.231% Params, 0.151 GMac, 2.805% MACs, 512, 512, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(1.049 M, 4.103% Params, 0.067 GMac, 1.247% MACs, 512, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(0.004 M, 0.016% Params, 0.0 GMac, 0.005% MACs, 2048, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.005% MACs, inplace=True)
(downsample): Sequential(
2.101 M, 8.222% Params, 0.134 GMac, 2.498% MACs,

```
            (0): Conv2d(2.097 M, 8.206% Params, 0.134 GMac, 2.493% MACs, 1024, 2048,
kernel_size=(1, 1), stride=(2, 2), bias=False)
            (1): BatchNorm2d(0.004 M, 0.016% Params, 0.0 GMac, 0.005% MACs, 2048, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
          )
        )
        (1): Bottleneck(
          4.463 M, 17.461% Params, 0.286 GMac, 5.309% MACs,
          (conv1): Conv2d(1.049 M, 4.103% Params, 0.067 GMac, 1.247% MACs, 2048, 512,
kernel_size=(1, 1), stride=(1, 1), bias=False)
          (bn1): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
          (conv2): Conv2d(2.359 M, 9.231% Params, 0.151 GMac, 2.805% MACs, 512, 512,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
          (bn2): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
          (conv3): Conv2d(1.049 M, 4.103% Params, 0.067 GMac, 1.247% MACs, 512, 2048,
kernel_size=(1, 1), stride=(1, 1), bias=False)
          (bn3): BatchNorm2d(0.004 M, 0.016% Params, 0.0 GMac, 0.005% MACs, 2048, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
          (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.004% MACs, inplace=True)
        )
        (2): Bottleneck(
          4.463 M, 17.461% Params, 0.286 GMac, 5.309% MACs,
          (conv1): Conv2d(1.049 M, 4.103% Params, 0.067 GMac, 1.247% MACs, 2048, 512,
kernel_size=(1, 1), stride=(1, 1), bias=False)
          (bn1): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
          (conv2): Conv2d(2.359 M, 9.231% Params, 0.151 GMac, 2.805% MACs, 512, 512,
kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
          (bn2): BatchNorm2d(0.001 M, 0.004% Params, 0.0 GMac, 0.001% MACs, 512, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
          (conv3): Conv2d(1.049 M, 4.103% Params, 0.067 GMac, 1.247% MACs, 512, 2048,
kernel_size=(1, 1), stride=(1, 1), bias=False)
          (bn3): BatchNorm2d(0.004 M, 0.016% Params, 0.0 GMac, 0.005% MACs, 2048, eps=1e-
05, momentum=0.1, affine=True, track_running_stats=True)
          (relu): ReLU(0.0 M, 0.000% Params, 0.0 GMac, 0.004% MACs, inplace=True)
        )
      )
      (avgpool): AdaptiveAvgPool2d(0.0 M, 0.000% Params, 0.0 GMac, 0.002% MACs,
output_size=(1, 1))
      (fc): Linear(2.049 M, 8.017% Params, 0.002 GMac, 0.038% MACs, in_features=2048,
out_features=1000, bias=True)
    )
```

Computational complexity:	5383111656.0
Number of parameters:	25557032