```
------------------------ DeepSpeed Flops Profiler ------------------------
Profile Summary at step 10:
Notations:
data parallel size (dp_size), model parallel size(mp_size),
number of parameters (params), number of multiply-accumulate operations(MACs),
number of floating-point operations (flops), floating-point operations per second
(FLOPS),
fwd latency (forward propagation latency), bwd latency (backward propagation
latency),
step (weights update latency), iter latency (sum of fwd, bwd and step latency)

params per gpu:                                         23.51 M
params of model = params per GPU * mp_size:            23.51 M
fwd MACs per GPU:                                       32.7 GMACs
fwd flops per GPU:                                      65.66 G
fwd flops of model = fwd flops per GPU * mp_size:      65.66 G
fwd latency:                                            95.94 ms
fwd FLOPS per GPU = fwd flops per GPU / fwd latency:   684.34 GFLOPS

--------------------------- Aggregated Profile per GPU
-----------------------------
Top 1 modules in terms of params, MACs or fwd latency at different model depths:
depth 0:
    params      - {'ResNetModel': '23.51 M'}
    MACs        - {'ResNetModel': '32.7 GMACs'}
    fwd latency - {'ResNetModel': '95.94 ms'}
depth 1:
    params      - {'ResNetEncoder': '23.5 M'}
    MACs        - {'ResNetEncoder': '31.75 GMACs'}
    fwd latency - {'ResNetEncoder': '93.53 ms'}
depth 2:
    params      - {'ModuleList': '23.5 M'}
    MACs        - {'ModuleList': '31.75 GMACs'}
    fwd latency - {'ModuleList': '93.37 ms'}
depth 3:
    params      - {'ResNetStage': '23.5 M'}
    MACs        - {'ResNetStage': '31.75 GMACs'}
    fwd latency - {'ResNetStage': '93.37 ms'}
depth 4:
    params      - {'Sequential': '23.5 M'}
    MACs        - {'Sequential': '31.75 GMACs'}
    fwd latency - {'Sequential': '92.89 ms'}
depth 5:
    params      - {'ResNetBottleNeckLayer': '23.5 M'}
    MACs        - {'ResNetBottleNeckLayer': '31.75 GMACs'}
    fwd latency - {'ResNetBottleNeckLayer': '92.89 ms'}
depth 6:
    params      - {'Sequential': '20.72 M'}
    MACs        - {'Sequential': '28.88 GMACs'}
```

```
    fwd latency - {'Sequential': '78.91 ms'}
depth 7:
    params        - {'ResNetConvLayer': '20.72 M'}
    MACs          - {'ResNetConvLayer': '28.88 GMACs'}
    fwd latency - {'ResNetConvLayer': '77.34 ms'}

---------------------------- Detailed Profile per GPU
----------------------------
Each module profile is listed after its name in the following order:
params, percentage of total params, MACs, percentage of total MACs, fwd latency,
percentage of total fwd latency, fwd FLOPS

Note: 1. A module can have torch.nn.module or torch.nn.functional to compute logits
(e.g. CrossEntropyLoss). They are not counted as submodules, thus not to be printed
out. However they make up the difference between a parent's MACs (or latency) and
the sum of its submodules'.
2. Number of floating-point operations is a theoretical estimation, thus FLOPS
computed using that could be larger than the maximum system throughput.
3. The fwd latency listed in the top module's profile is directly captured at the
module forward function in PyTorch, thus it's less than the fwd latency shown above
which is captured in DeepSpeed.

ResNetModel(
  23.51 M, 100.00% Params, 32.7 GMACs, 100.00% MACs, 95.94 ms, 100.00% latency,
684.34 GFLOPS,
  (embedder): ResNetEmbeddings(
    9.54 k, 0.04% Params, 944.11 MMACs, 2.89% MACs, 1.88 ms, 1.96% latency, 1.02
TFLOPS,
    (embedder): ResNetConvLayer(
      9.54 k, 0.04% Params, 944.11 MMACs, 2.89% MACs, 1.64 ms, 1.71% latency, 1.17
TFLOPS,
      (convolution): Conv2d(9.41 k, 0.04% Params, 944.11 MMACs, 2.89% MACs, 180.01
us, 0.19% latency, 10.49 TFLOPS, 3, 64, kernel_size=(7, 7), stride=(2, 2),
padding=(3, 3), bias=False)
      (normalization): BatchNorm2d(128, 0.00% Params, 0 MACs, 0.00% MACs, 958.2 us,
1.00% latency, 13.41 GFLOPS, 64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 396.49 us, 0.41%
latency, 16.2 GFLOPS, )
    )
    (pooler): MaxPool2d(0, 0.00% Params, 0 MACs, 0.00% MACs, 165.22 us, 0.17%
latency, 38.87 GFLOPS, kernel_size=3, stride=2, padding=1, dilation=1,
ceil_mode=False)
  )
  (encoder): ResNetEncoder(
    23.5 M, 99.96% Params, 31.75 GMACs, 97.11% MACs, 93.53 ms, 97.49% latency,
681.49 GFLOPS,
    (stages): ModuleList(
      (0): ResNetStage(
        215.81 k, 0.92% Params, 5.34 GMACs, 16.34% MACs, 15.51 ms, 16.16% latency,
```

695.59 GFLOPS,
        (layers): Sequential(
          215.81 k, 0.92% Params, 5.34 GMACs, 16.34% MACs, 15.42 ms, 16.08% latency,
699.31 GFLOPS,
          (0): ResNetBottleNeckLayer(
            75.01 k, 0.32% Params, 1.85 GMACs, 5.66% MACs, 5.92 ms, 6.17% latency,
631.57 GFLOPS,
            (shortcut): ResNetShortCut(
              16.9 k, 0.07% Params, 411.04 MMACs, 1.26% MACs, 1.1 ms, 1.15% latency,
759.81 GFLOPS,
              (convolution): Conv2d(16.38 k, 0.07% Params, 411.04 MMACs, 1.26% MACs,
141.62 us, 0.15% latency, 5.8 TFLOPS, 64, 256, kernel_size=(1, 1), stride=(1, 1),
bias=False)
              (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
891.92 us, 0.93% latency, 14.4 GFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
)
            (layer): Sequential(
              58.11 k, 0.25% Params, 1.44 GMACs, 4.40% MACs, 4.32 ms, 4.50% latency,
671.85 GFLOPS,
              (0): ResNetConvLayer(
                4.22 k, 0.02% Params, 102.76 MMACs, 0.31% MACs, 1.55 ms, 1.61%
latency, 135.87 GFLOPS,
                (convolution): Conv2d(4.1 k, 0.02% Params, 102.76 MMACs, 0.31% MACs,
140.43 us, 0.15% latency, 1.46 TFLOPS, 64, 64, kernel_size=(1, 1), stride=(1, 1),
bias=False)
                (normalization): BatchNorm2d(128, 0.00% Params, 0 MACs, 0.00% MACs,
938.65 us, 0.98% latency, 3.42 GFLOPS, 64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 374.32 us,
0.39% latency, 4.29 GFLOPS,)
)
              (1): ResNetConvLayer(
                36.99 k, 0.16% Params, 924.84 MMACs, 2.83% MACs, 1.53 ms, 1.59%
latency, 1.21 TFLOPS,
                (convolution): Conv2d(36.86 k, 0.16% Params, 924.84 MMACs, 2.83%
MACs, 137.81 us, 0.14% latency, 13.42 TFLOPS, 64, 64, kernel_size=(3, 3), stride=(1,
1), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(128, 0.00% Params, 0 MACs, 0.00% MACs,
888.35 us, 0.93% latency, 3.61 GFLOPS, 64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 406.74 us,
0.42% latency, 3.95 GFLOPS,)
)
              (2): ResNetConvLayer(
                16.9 k, 0.07% Params, 411.04 MMACs, 1.26% MACs, 1.16 ms, 1.21%
latency, 721.16 GFLOPS,
                (convolution): Conv2d(16.38 k, 0.07% Params, 411.04 MMACs, 1.26%
MACs, 138.04 us, 0.14% latency, 5.96 TFLOPS, 64, 256, kernel_size=(1, 1), stride=(1,
1), bias=False)

(normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
906.47 us, 0.94% latency, 14.17 GFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.17
us, 0.02% latency, 0.0 FLOPS,)
)
)
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 383.62 us, 0.40%
latency, 16.74 GFLOPS,)
)
          (1): ResNetBottleNeckLayer(
            70.4 k, 0.30% Params, 1.75 GMACs, 5.34% MACs, 4.75 ms, 4.95% latency,
741.56 GFLOPS,
            (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 17.4 us, 0.02%
latency, 0.0 FLOPS,)
            (layer): Sequential(
              70.4 k, 0.30% Params, 1.75 GMACs, 5.34% MACs, 4.25 ms, 4.43% latency,
826.71 GFLOPS,
              (0): ResNetConvLayer(
                16.51 k, 0.07% Params, 411.04 MMACs, 1.26% MACs, 1.53 ms, 1.59%
latency, 542.09 GFLOPS,
                (convolution): Conv2d(16.38 k, 0.07% Params, 411.04 MMACs, 1.26%
MACs, 145.44 us, 0.15% latency, 5.65 TFLOPS, 256, 64, kernel_size=(1, 1), stride=(1,
1), bias=False)
                (normalization): BatchNorm2d(128, 0.00% Params, 0 MACs, 0.00% MACs,
920.3 us, 0.96% latency, 3.49 GFLOPS, 64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 368.83 us,
0.38% latency, 4.35 GFLOPS,)
)
              (1): ResNetConvLayer(
                36.99 k, 0.16% Params, 924.84 MMACs, 2.83% MACs, 1.51 ms, 1.57%
latency, 1.23 TFLOPS,
                (convolution): Conv2d(36.86 k, 0.16% Params, 924.84 MMACs, 2.83%
MACs, 134.47 us, 0.14% latency, 13.76 TFLOPS, 64, 64, kernel_size=(3, 3), stride=(1,
1), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(128, 0.00% Params, 0 MACs, 0.00% MACs,
879.76 us, 0.92% latency, 3.65 GFLOPS, 64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 398.87 us,
0.42% latency, 4.03 GFLOPS,)
)
              (2): ResNetConvLayer(
                16.9 k, 0.07% Params, 411.04 MMACs, 1.26% MACs, 1.14 ms, 1.19%
latency, 732.93 GFLOPS,
                (convolution): Conv2d(16.38 k, 0.07% Params, 411.04 MMACs, 1.26%
MACs, 136.85 us, 0.14% latency, 6.01 TFLOPS, 64, 256, kernel_size=(1, 1), stride=(1,
1), bias=False)
                (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
894.07 us, 0.93% latency, 14.37 GFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,

```
track_running_stats=True)
              (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 20.5 us,
0.02% latency, 0.0 FLOPS, )
            )
          )
          (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 378.37 us, 0.39%
latency, 16.97 GFLOPS, )
        )
        (2): ResNetBottleNeckLayer(
          70.4 k, 0.30% Params, 1.75 GMACs, 5.34% MACs, 4.75 ms, 4.95% latency,
741.52 GFLOPS,
          (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 17.88 us,
0.02% latency, 0.0 FLOPS, )
          (layer): Sequential(
            70.4 k, 0.30% Params, 1.75 GMACs, 5.34% MACs, 4.24 ms, 4.42% latency,
828.9 GFLOPS,
            (0): ResNetConvLayer(
              16.51 k, 0.07% Params, 411.04 MMACs, 1.26% MACs, 1.5 ms, 1.57%
latency, 549.82 GFLOPS,
              (convolution): Conv2d(16.38 k, 0.07% Params, 411.04 MMACs, 1.26%
MACs, 160.22 us, 0.17% latency, 5.13 TFLOPS, 256, 64, kernel_size=(1, 1), stride=(1,
1), bias=False)
              (normalization): BatchNorm2d(128, 0.00% Params, 0 MACs, 0.00% MACs,
851.39 us, 0.89% latency, 3.77 GFLOPS, 64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
              (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 397.92 us,
0.41% latency, 4.04 GFLOPS, )
            )
            (1): ResNetConvLayer(
              36.99 k, 0.16% Params, 924.84 MMACs, 2.83% MACs, 1.48 ms, 1.54%
latency, 1.26 TFLOPS,
              (convolution): Conv2d(36.86 k, 0.16% Params, 924.84 MMACs, 2.83%
MACs, 134.71 us, 0.14% latency, 13.73 TFLOPS, 64, 64, kernel_size=(3, 3), stride=(1,
1), padding=(1, 1), bias=False)
              (normalization): BatchNorm2d(128, 0.00% Params, 0 MACs, 0.00% MACs,
879.53 us, 0.92% latency, 3.65 GFLOPS, 64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
              (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 366.45 us,
0.38% latency, 4.38 GFLOPS, )
            )
            (2): ResNetConvLayer(
              16.9 k, 0.07% Params, 411.04 MMACs, 1.26% MACs, 1.18 ms, 1.23%
latency, 705.75 GFLOPS,
              (convolution): Conv2d(16.38 k, 0.07% Params, 411.04 MMACs, 1.26%
MACs, 184.54 us, 0.19% latency, 4.45 TFLOPS, 64, 256, kernel_size=(1, 1), stride=(1,
1), bias=False)
              (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
863.79 us, 0.90% latency, 14.87 GFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
              (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.17
```

us, 0.02% latency, 0.0 FLOPS,)
)
)
              (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 386.0 us, 0.40%
latency, 16.64 GFLOPS,)
)
)
)
      (1): ResNetStage(
        1.22 M, 5.19% Params, 8.22 GMACs, 25.14% MACs, 21.17 ms, 22.07% latency,
779.91 GFLOPS,
        (layers): Sequential(
          1.22 M, 5.19% Params, 8.22 GMACs, 25.14% MACs, 21.07 ms, 21.96% latency,
783.87 GFLOPS,
          (0): ResNetBottleNeckLayer(
            379.39 k, 1.61% Params, 2.98 GMACs, 9.11% MACs, 5.89 ms, 6.14% latency,
1.02 TFLOPS,
            (shortcut): ResNetShortCut(
              132.1 k, 0.56% Params, 822.08 MMACs, 2.51% MACs, 1.14 ms, 1.19%
latency, 1.45 TFLOPS,
              (convolution): Conv2d(131.07 k, 0.56% Params, 822.08 MMACs, 2.51%
MACs, 180.01 us, 0.19% latency, 9.13 TFLOPS, 256, 512, kernel_size=(1, 1),
stride=(2, 2), bias=False)
              (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00% MACs,
890.25 us, 0.93% latency, 7.21 GFLOPS, 512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
)
            (layer): Sequential(
              247.3 k, 1.05% Params, 2.16 GMACs, 6.60% MACs, 4.26 ms, 4.44% latency,
1.02 TFLOPS,
              (0): ResNetConvLayer(
                33.02 k, 0.14% Params, 822.08 MMACs, 2.51% MACs, 1.52 ms, 1.58%
latency, 1.09 TFLOPS,
                (convolution): Conv2d(32.77 k, 0.14% Params, 822.08 MMACs, 2.51%
MACs, 139.47 us, 0.15% latency, 11.79 TFLOPS, 256, 128, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(256, 0.00% Params, 0 MACs, 0.00% MACs,
882.39 us, 0.92% latency, 7.28 GFLOPS, 128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 401.5 us,
0.42% latency, 8.0 GFLOPS,)
)
              (1): ResNetConvLayer(
                147.71 k, 0.63% Params, 924.84 MMACs, 2.83% MACs, 1.49 ms, 1.55%
latency, 1.24 TFLOPS,
                (convolution): Conv2d(147.46 k, 0.63% Params, 924.84 MMACs, 2.83%
MACs, 135.18 us, 0.14% latency, 13.68 TFLOPS, 128, 128, kernel_size=(3, 3),
stride=(2, 2), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(256, 0.00% Params, 0 MACs, 0.00% MACs,
879.76 us, 0.92% latency, 1.83 GFLOPS, 128, eps=1e-05, momentum=0.1, affine=True,

track_running_stats=True)
                    (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 379.32 us,
0.40% latency, 2.12 GFLOPS,)
)
                (2): ResNetConvLayer(
                    66.56 k, 0.28% Params, 411.04 MMACs, 1.26% MACs, 1.18 ms, 1.23%
latency, 702.59 GFLOPS,
                    (convolution): Conv2d(65.54 k, 0.28% Params, 411.04 MMACs, 1.26%
MACs, 135.42 us, 0.14% latency, 6.07 TFLOPS, 128, 512, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                    (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 930.07 us, 0.97% latency, 6.91 GFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                    (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.41
us, 0.02% latency, 0.0 FLOPS,)
)
)
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 380.75 us, 0.40%
latency, 8.43 GFLOPS,)
)
        (1): ResNetBottleNeckLayer(
            280.06 k, 1.19% Params, 1.75 GMACs, 5.34% MACs, 5.26 ms, 5.48% latency,
666.89 GFLOPS,
            (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 17.4 us, 0.02%
latency, 0.0 FLOPS,)
            (layer): Sequential(
                280.06 k, 1.19% Params, 1.75 GMACs, 5.34% MACs, 4.76 ms, 4.96%
latency, 736.07 GFLOPS,
                (0): ResNetConvLayer(
                    65.79 k, 0.28% Params, 411.04 MMACs, 1.26% MACs, 1.83 ms, 1.90%
latency, 451.34 GFLOPS,
                    (convolution): Conv2d(65.54 k, 0.28% Params, 411.04 MMACs, 1.26%
MACs, 304.22 us, 0.32% latency, 2.7 TFLOPS, 512, 128, kernel_size=(1, 1), stride=(1,
1), bias=False)
                    (normalization): BatchNorm2d(256, 0.00% Params, 0 MACs, 0.00% MACs,
981.09 us, 1.02% latency, 1.64 GFLOPS, 128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                    (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 431.78 us,
0.45% latency, 1.86 GFLOPS,)
)
                (1): ResNetConvLayer(
                    147.71 k, 0.63% Params, 924.84 MMACs, 2.83% MACs, 1.65 ms, 1.72%
latency, 1.12 TFLOPS,
                    (convolution): Conv2d(147.46 k, 0.63% Params, 924.84 MMACs, 2.83%
MACs, 136.61 us, 0.14% latency, 13.54 TFLOPS, 128, 128, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
                    (normalization): BatchNorm2d(256, 0.00% Params, 0 MACs, 0.00% MACs,
970.6 us, 1.01% latency, 1.65 GFLOPS, 128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                    (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 436.54 us,

```
0.46% latency, 1.84 GFLOPS, )
              )
              (2): ResNetConvLayer(
                66.56 k, 0.28% Params, 411.04 MMACs, 1.26% MACs, 1.2 ms, 1.25%
latency, 689.62 GFLOPS,
                (convolution): Conv2d(65.54 k, 0.28% Params, 411.04 MMACs, 1.26%
MACs, 155.69 us, 0.16% latency, 5.28 TFLOPS, 128, 512, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 935.32 us, 0.97% latency, 6.87 GFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 21.22
us, 0.02% latency, 0.0 FLOPS, )
              )
            )
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 379.8 us, 0.40%
latency, 8.46 GFLOPS, )
          )
          (2): ResNetBottleNeckLayer(
            280.06 k, 1.19% Params, 1.75 GMACs, 5.34% MACs, 4.83 ms, 5.03% latency,
726.34 GFLOPS,
            (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 18.36 us,
0.02% latency, 0.0 FLOPS, )
            (layer): Sequential(
              280.06 k, 1.19% Params, 1.75 GMACs, 5.34% MACs, 4.24 ms, 4.42%
latency, 826.53 GFLOPS,
              (0): ResNetConvLayer(
                65.79 k, 0.28% Params, 411.04 MMACs, 1.26% MACs, 1.52 ms, 1.58%
latency, 543.05 GFLOPS,
                (convolution): Conv2d(65.54 k, 0.28% Params, 411.04 MMACs, 1.26%
MACs, 164.03 us, 0.17% latency, 5.01 TFLOPS, 512, 128, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(256, 0.00% Params, 0 MACs, 0.00% MACs,
853.06 us, 0.89% latency, 1.88 GFLOPS, 128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 406.27 us,
0.42% latency, 1.98 GFLOPS, )
              )
              (1): ResNetConvLayer(
                147.71 k, 0.63% Params, 924.84 MMACs, 2.83% MACs, 1.49 ms, 1.55%
latency, 1.25 TFLOPS,
                (convolution): Conv2d(147.46 k, 0.63% Params, 924.84 MMACs, 2.83%
MACs, 135.9 us, 0.14% latency, 13.61 TFLOPS, 128, 128, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(256, 0.00% Params, 0 MACs, 0.00% MACs,
879.76 us, 0.92% latency, 1.83 GFLOPS, 128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 371.69 us,
0.39% latency, 2.16 GFLOPS, )
              )
```

```
            (2): ResNetConvLayer(
              66.56 k, 0.28% Params, 411.04 MMACs, 1.26% MACs, 1.16 ms, 1.21%
latency, 714.87 GFLOPS,
              (convolution): Conv2d(65.54 k, 0.28% Params, 411.04 MMACs, 1.26%
MACs, 166.18 us, 0.17% latency, 4.95 TFLOPS, 128, 512, kernel_size=(1, 1),
stride=(1, 1), bias=False)
              (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 858.78 us, 0.90% latency, 7.48 GFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
              (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 21.22
us, 0.02% latency, 0.0 FLOPS, )
            )
          )
          (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 428.2 us, 0.45%
latency, 7.5 GFLOPS, )
        )
        (3): ResNetBottleNeckLayer(
          280.06 k, 1.19% Params, 1.75 GMACs, 5.34% MACs, 5.09 ms, 5.30% latency,
689.77 GFLOPS,
          (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 17.4 us, 0.02%
latency, 0.0 FLOPS, )
          (layer): Sequential(
            280.06 k, 1.19% Params, 1.75 GMACs, 5.34% MACs, 4.58 ms, 4.78%
latency, 764.51 GFLOPS,
            (0): ResNetConvLayer(
              65.79 k, 0.28% Params, 411.04 MMACs, 1.26% MACs, 1.69 ms, 1.77%
latency, 486.86 GFLOPS,
              (convolution): Conv2d(65.54 k, 0.28% Params, 411.04 MMACs, 1.26%
MACs, 154.02 us, 0.16% latency, 5.34 TFLOPS, 512, 128, kernel_size=(1, 1),
stride=(1, 1), bias=False)
              (normalization): BatchNorm2d(256, 0.00% Params, 0 MACs, 0.00% MACs,
1.01 ms, 1.05% latency, 1.59 GFLOPS, 128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
              (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 429.87 us,
0.45% latency, 1.87 GFLOPS, )
            )
            (1): ResNetConvLayer(
              147.71 k, 0.63% Params, 924.84 MMACs, 2.83% MACs, 1.6 ms, 1.67%
latency, 1.16 TFLOPS,
              (convolution): Conv2d(147.46 k, 0.63% Params, 924.84 MMACs, 2.83%
MACs, 143.29 us, 0.15% latency, 12.91 TFLOPS, 128, 128, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
              (normalization): BatchNorm2d(256, 0.00% Params, 0 MACs, 0.00% MACs,
933.41 us, 0.97% latency, 1.72 GFLOPS, 128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
              (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 405.07 us,
0.42% latency, 1.98 GFLOPS, )
            )
            (2): ResNetConvLayer(
              66.56 k, 0.28% Params, 411.04 MMACs, 1.26% MACs, 1.2 ms, 1.25%
```

latency, 692.23 GFLOPS,
                (convolution): Conv2d(65.54 k, 0.28% Params, 411.04 MMACs, 1.26%
MACs, 149.25 us, 0.16% latency, 5.51 TFLOPS, 128, 512, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 930.31 us, 0.97% latency, 6.9 GFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.17
us, 0.02% latency, 0.0 FLOPS,)
)
)
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 379.32 us, 0.40%
latency, 8.47 GFLOPS,)
)
)
)
      (2): ResNetStage(
        7.1 M, 30.20% Params, 11.71 GMACs, 35.83% MACs, 37.25 ms, 38.83% latency,
630.26 GFLOPS,
        (layers): Sequential(
          7.1 M, 30.20% Params, 11.71 GMACs, 35.83% MACs, 37.06 ms, 38.63% latency,
633.5 GFLOPS,
          (0): ResNetBottleNeckLayer(
            1.51 M, 6.43% Params, 2.98 GMACs, 9.11% MACs, 7.09 ms, 7.39% latency,
842.86 GFLOPS,
            (shortcut): ResNetShortCut(
              526.34 k, 2.24% Params, 822.08 MMACs, 2.51% MACs, 1.26 ms, 1.31%
latency, 1.31 TFLOPS,
              (convolution): Conv2d(524.29 k, 2.23% Params, 822.08 MMACs, 2.51%
MACs, 178.81 us, 0.19% latency, 9.19 TFLOPS, 512, 1024, kernel_size=(1, 1),
stride=(2, 2), bias=False)
              (normalization): BatchNorm2d(2.05 k, 0.01% Params, 0 MACs, 0.00% MACs,
1.0 ms, 1.04% latency, 3.2 GFLOPS, 1024, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
)
            (layer): Sequential(
              986.11 k, 4.19% Params, 2.16 GMACs, 6.60% MACs, 5.26 ms, 5.48%
latency, 822.76 GFLOPS,
              (0): ResNetConvLayer(
                131.58 k, 0.56% Params, 822.08 MMACs, 2.51% MACs, 1.92 ms, 2.00%
latency, 858.85 GFLOPS,
                (convolution): Conv2d(131.07 k, 0.56% Params, 822.08 MMACs, 2.51%
MACs, 165.22 us, 0.17% latency, 9.95 TFLOPS, 512, 256, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
1.01 ms, 1.05% latency, 3.18 GFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 617.74 us,
0.64% latency, 2.6 GFLOPS,)
)

```
        (1): ResNetConvLayer(
          590.34 k, 2.51% Params, 924.84 MMACs, 2.83% MACs, 1.93 ms, 2.02%
latency, 957.36 GFLOPS,
          (convolution): Conv2d(589.82 k, 2.51% Params, 924.84 MMACs, 2.83%
MACs, 226.97 us, 0.24% latency, 8.15 TFLOPS, 256, 256, kernel_size=(3, 3),
stride=(2, 2), padding=(1, 1), bias=False)
          (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
970.36 us, 1.01% latency, 827.34 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
          (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 617.27 us,
0.64% latency, 650.3 MFLOPS, )
        )
        (2): ResNetConvLayer(
          264.19 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.28 ms, 1.34%
latency, 642.33 GFLOPS,
          (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 194.55 us, 0.20% latency, 4.23 TFLOPS, 256, 1024, kernel_size=(1, 1),
stride=(1, 1), bias=False)
          (normalization): BatchNorm2d(2.05 k, 0.01% Params, 0 MACs, 0.00%
MACs, 956.06 us, 1.00% latency, 3.36 GFLOPS, 1024, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 23.37
us, 0.02% latency, 0.0 FLOPS, )
        )
      )
      (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 457.53 us, 0.48%
latency, 3.51 GFLOPS, )
    )
    (1): ResNetBottleNeckLayer(
      1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 5.21 ms, 5.43% latency,
672.49 GFLOPS,
      (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 18.6 us, 0.02%
latency, 0.0 FLOPS, )
      (layer): Sequential(
        1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 4.62 ms, 4.82% latency,
756.86 GFLOPS,
        (0): ResNetConvLayer(
          262.66 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.63 ms, 1.70%
latency, 503.59 GFLOPS,
          (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 156.16 us, 0.16% latency, 5.26 TFLOPS, 1024, 256, kernel_size=(1, 1),
stride=(1, 1), bias=False)
          (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
955.58 us, 1.00% latency, 840.13 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
          (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 419.62 us,
0.44% latency, 956.61 MFLOPS, )
        )
        (1): ResNetConvLayer(
          590.34 k, 2.51% Params, 924.84 MMACs, 2.83% MACs, 1.67 ms, 1.74%
```

```
latency, 1.11 TFLOPS,
                (convolution): Conv2d(589.82 k, 2.51% Params, 924.84 MMACs, 2.83%
MACs, 151.4 us, 0.16% latency, 12.22 TFLOPS, 256, 256, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
988.25 us, 1.03% latency, 812.37 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 426.53 us,
0.44% latency, 941.1 MFLOPS, )
              )
            (2): ResNetConvLayer(
              264.19 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.2 ms, 1.25%
latency, 688.45 GFLOPS,
                (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 144.96 us, 0.15% latency, 5.67 TFLOPS, 256, 1024, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(2.05 k, 0.01% Params, 0 MACs, 0.00%
MACs, 936.51 us, 0.98% latency, 3.43 GFLOPS, 1024, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.89
us, 0.02% latency, 0.0 FLOPS, )
              )
            )
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 431.06 us, 0.45%
latency, 3.72 GFLOPS, )
          )
          (2): ResNetBottleNeckLayer(
            1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 5.03 ms, 5.24% latency,
696.22 GFLOPS,
            (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 19.31 us,
0.02% latency, 0.0 FLOPS, )
            (layer): Sequential(
              1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 4.47 ms, 4.66% latency,
783.2 GFLOPS,
              (0): ResNetConvLayer(
                262.66 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.56 ms, 1.63%
latency, 526.47 GFLOPS,
                (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 146.15 us, 0.15% latency, 5.62 TFLOPS, 1024, 256, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
928.4 us, 0.97% latency, 864.73 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 390.77 us,
0.41% latency, 1.03 GFLOPS, )
              )
              (1): ResNetConvLayer(
                590.34 k, 2.51% Params, 924.84 MMACs, 2.83% MACs, 1.62 ms, 1.69%
latency, 1.14 TFLOPS,
                (convolution): Conv2d(589.82 k, 2.51% Params, 924.84 MMACs, 2.83%
```

MACs, 142.57 us, 0.15% latency, 12.97 TFLOPS, 256, 256, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
920.53 us, 0.96% latency, 872.12 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 444.41 us,
0.46% latency, 903.23 MFLOPS,)
)
              (2): ResNetConvLayer(
                264.19 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.2 ms, 1.25%
latency, 688.45 GFLOPS,
                (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 154.73 us, 0.16% latency, 5.31 TFLOPS, 256, 1024, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(2.05 k, 0.01% Params, 0 MACs, 0.00%
MACs, 924.83 us, 0.96% latency, 3.47 GFLOPS, 1024, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.41
us, 0.02% latency, 0.0 FLOPS,)
)
)
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 432.97 us, 0.45%
latency, 3.71 GFLOPS,)
)
          (3): ResNetBottleNeckLayer(
            1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 5.03 ms, 5.24% latency,
696.25 GFLOPS,
            (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 18.6 us, 0.02%
latency, 0.0 FLOPS,)
            (layer): Sequential(
              1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 4.47 ms, 4.66% latency,
783.11 GFLOPS,
              (0): ResNetConvLayer(
                262.66 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.58 ms, 1.65%
latency, 521.38 GFLOPS,
                (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 145.44 us, 0.15% latency, 5.65 TFLOPS, 1024, 256, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
945.81 us, 0.99% latency, 848.82 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 387.67 us,
0.40% latency, 1.04 GFLOPS,)
)
              (1): ResNetConvLayer(
                590.34 k, 2.51% Params, 924.84 MMACs, 2.83% MACs, 1.61 ms, 1.68%
latency, 1.15 TFLOPS,
                (convolution): Conv2d(589.82 k, 2.51% Params, 924.84 MMACs, 2.83%
MACs, 178.1 us, 0.19% latency, 10.39 TFLOPS, 256, 256, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)

(normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
927.69 us, 0.97% latency, 865.4 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                    (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 401.02 us,
0.42% latency, 1.0 GFLOPS,)
)
                  (2): ResNetConvLayer(
                    264.19 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.19 ms, 1.24%
latency, 693.69 GFLOPS,
                    (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 142.81 us, 0.15% latency, 5.76 TFLOPS, 256, 1024, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                    (normalization): BatchNorm2d(2.05 k, 0.01% Params, 0 MACs, 0.00%
MACs, 930.07 us, 0.97% latency, 3.45 GFLOPS, 1024, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                    (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.17
us, 0.02% latency, 0.0 FLOPS,)
)
)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 433.92 us, 0.45%
latency, 3.7 GFLOPS,)
)
              (4): ResNetBottleNeckLayer(
                1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 7.51 ms, 7.83% latency,
466.25 GFLOPS,
                (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 29.8 us, 0.03%
latency, 0.0 FLOPS,)
                (layer): Sequential(
                  1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 6.74 ms, 7.03% latency,
518.87 GFLOPS,
                  (0): ResNetConvLayer(
                    262.66 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 2.12 ms, 2.21%
latency, 388.08 GFLOPS,
                    (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 144.24 us, 0.15% latency, 5.7 TFLOPS, 1024, 256, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                    (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
1.17 ms, 1.22% latency, 683.71 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                    (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 682.59 us,
0.71% latency, 588.06 MFLOPS,)
)
                  (1): ResNetConvLayer(
                    590.34 k, 2.51% Params, 924.84 MMACs, 2.83% MACs, 2.6 ms, 2.71%
latency, 711.57 GFLOPS,
                    (convolution): Conv2d(589.82 k, 2.51% Params, 924.84 MMACs, 2.83%
MACs, 269.89 us, 0.28% latency, 6.85 TFLOPS, 256, 256, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
                    (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
1.66 ms, 1.73% latency, 482.83 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,

```
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 437.26 us,
0.46% latency, 918.01 MFLOPS, )
              )
              (2): ResNetConvLayer(
                264.19 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.86 ms, 1.94%
latency, 442.6 GFLOPS,
                (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 247.96 us, 0.26% latency, 3.32 TFLOPS, 256, 1024, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(2.05 k, 0.01% Params, 0 MACs, 0.00%
MACs, 1.44 ms, 1.50% latency, 2.22 GFLOPS, 1024, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 32.9 us,
0.03% latency, 0.0 FLOPS, )
              )
            )
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 581.74 us, 0.61%
latency, 2.76 GFLOPS, )
          )
          (5): ResNetBottleNeckLayer(
            1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 7.2 ms, 7.51% latency,
486.07 GFLOPS,
            (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 19.07 us,
0.02% latency, 0.0 FLOPS, )
            (layer): Sequential(
              1.12 M, 4.75% Params, 1.75 GMACs, 5.34% MACs, 6.37 ms, 6.64% latency,
549.03 GFLOPS,
              (0): ResNetConvLayer(
                262.66 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 2.58 ms, 2.69%
latency, 319.11 GFLOPS,
                (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 233.65 us, 0.24% latency, 3.52 TFLOPS, 1024, 256, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
1.34 ms, 1.39% latency, 599.9 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 864.27 us,
0.90% latency, 464.45 MFLOPS, )
              )
              (1): ResNetConvLayer(
                590.34 k, 2.51% Params, 924.84 MMACs, 2.83% MACs, 2.36 ms, 2.45%
latency, 785.91 GFLOPS,
                (convolution): Conv2d(589.82 k, 2.51% Params, 924.84 MMACs, 2.83%
MACs, 235.56 us, 0.25% latency, 7.85 TFLOPS, 256, 256, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(512, 0.00% Params, 0 MACs, 0.00% MACs,
1.54 ms, 1.61% latency, 520.44 MFLOPS, 256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 432.73 us,
```

```
0.45% latency, 927.62 MFLOPS, )
              )
              (2): ResNetConvLayer(
                264.19 k, 1.12% Params, 411.04 MMACs, 1.26% MACs, 1.33 ms, 1.39%
latency, 618.46 GFLOPS,
                (convolution): Conv2d(262.14 k, 1.12% Params, 411.04 MMACs, 1.26%
MACs, 175.71 us, 0.18% latency, 4.68 TFLOPS, 256, 1024, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(2.05 k, 0.01% Params, 0 MACs, 0.00%
MACs, 1.01 ms, 1.06% latency, 3.17 GFLOPS, 1024, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.89
us, 0.02% latency, 0.0 FLOPS, )
              )
            )
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 642.54 us, 0.67%
latency, 2.5 GFLOPS, )
          )
        )
      )
      (3): ResNetStage(
        14.96 M, 63.66% Params, 6.47 GMACs, 19.80% MACs, 19.44 ms, 20.26% latency,
666.76 GFLOPS,
        (layers): Sequential(
          14.96 M, 63.66% Params, 6.47 GMACs, 19.80% MACs, 19.33 ms, 20.15% latency,
670.46 GFLOPS,
          (0): ResNetBottleNeckLayer(
            6.04 M, 25.69% Params, 2.98 GMACs, 9.11% MACs, 8.0 ms, 8.34% latency,
745.72 GFLOPS,
            (shortcut): ResNetShortCut(
              2.1 M, 8.94% Params, 822.08 MMACs, 2.51% MACs, 1.23 ms, 1.28% latency,
1.34 TFLOPS,
              (convolution): Conv2d(2.1 M, 8.92% Params, 822.08 MMACs, 2.51% MACs,
152.35 us, 0.16% latency, 10.79 TFLOPS, 1024, 2048, kernel_size=(1, 1), stride=(2,
2), bias=False)
              (normalization): BatchNorm2d(4.1 k, 0.02% Params, 0 MACs, 0.00% MACs,
1.01 ms, 1.05% latency, 1.59 GFLOPS, 2048, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
            )
            (layer): Sequential(
              3.94 M, 16.75% Params, 2.16 GMACs, 6.60% MACs, 6.16 ms, 6.42% latency,
701.17 GFLOPS,
              (0): ResNetConvLayer(
                525.31 k, 2.23% Params, 822.08 MMACs, 2.51% MACs, 2.95 ms, 3.08%
latency, 557.22 GFLOPS,
                (convolution): Conv2d(524.29 k, 2.23% Params, 822.08 MMACs, 2.51%
MACs, 260.11 us, 0.27% latency, 6.32 TFLOPS, 1024, 512, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 1.8 ms, 1.87% latency, 892.7 MFLOPS, 512, eps=1e-05, momentum=0.1,
```

affine=True, track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 730.75 us,
0.76% latency, 1.1 GFLOPS,)
)
              (1): ResNetConvLayer(
                2.36 M, 10.04% Params, 924.84 MMACs, 2.83% MACs, 1.8 ms, 1.88%
latency, 1.03 TFLOPS,
                (convolution): Conv2d(2.36 M, 10.04% Params, 924.84 MMACs, 2.83%
MACs, 188.11 us, 0.20% latency, 9.83 TFLOPS, 512, 512, kernel_size=(3, 3),
stride=(2, 2), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 1.05 ms, 1.10% latency, 380.57 MFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 431.78 us,
0.45% latency, 464.83 MFLOPS,)
)
              (2): ResNetConvLayer(
                1.05 M, 4.48% Params, 411.04 MMACs, 1.26% MACs, 1.26 ms, 1.32%
latency, 651.48 GFLOPS,
                (convolution): Conv2d(1.05 M, 4.46% Params, 411.04 MMACs, 1.26%
MACs, 161.89 us, 0.17% latency, 5.08 TFLOPS, 512, 2048, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(4.1 k, 0.02% Params, 0 MACs, 0.00%
MACs, 982.05 us, 1.02% latency, 1.63 GFLOPS, 2048, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.41
us, 0.02% latency, 0.0 FLOPS,)
)
)
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 476.12 us, 0.50%
latency, 1.69 GFLOPS,)
)
          (1): ResNetBottleNeckLayer(
            4.46 M, 18.98% Params, 1.75 GMACs, 5.34% MACs, 5.33 ms, 5.56% latency,
655.85 GFLOPS,
            (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 41.72 us,
0.04% latency, 0.0 FLOPS,)
            (layer): Sequential(
              4.46 M, 18.98% Params, 1.75 GMACs, 5.34% MACs, 4.76 ms, 4.96% latency,
734.22 GFLOPS,
              (0): ResNetConvLayer(
                1.05 M, 4.46% Params, 411.04 MMACs, 1.26% MACs, 1.69 ms, 1.76%
latency, 486.89 GFLOPS,
                (convolution): Conv2d(1.05 M, 4.46% Params, 411.04 MMACs, 1.26%
MACs, 162.36 us, 0.17% latency, 5.06 TFLOPS, 2048, 512, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 1.03 ms, 1.08% latency, 389.19 MFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 394.34 us,

```
0.41% latency, 508.96 MFLOPS, )
              )
            (1): ResNetConvLayer(
              2.36 M, 10.04% Params, 924.84 MMACs, 2.83% MACs, 1.79 ms, 1.86%
latency, 1.03 TFLOPS,
                (convolution): Conv2d(2.36 M, 10.04% Params, 924.84 MMACs, 2.83%
MACs, 212.19 us, 0.22% latency, 8.72 TFLOPS, 512, 512, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
                (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 1.03 ms, 1.08% latency, 388.2 MFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 421.05 us,
0.44% latency, 476.68 MFLOPS, )
              )
            (2): ResNetConvLayer(
              1.05 M, 4.48% Params, 411.04 MMACs, 1.26% MACs, 1.2 ms, 1.25%
latency, 687.93 GFLOPS,
                (convolution): Conv2d(1.05 M, 4.46% Params, 411.04 MMACs, 1.26%
MACs, 143.29 us, 0.15% latency, 5.74 TFLOPS, 512, 2048, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(4.1 k, 0.02% Params, 0 MACs, 0.00%
MACs, 934.12 us, 0.97% latency, 1.72 GFLOPS, 2048, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 23.13
us, 0.02% latency, 0.0 FLOPS, )
              )
            )
            (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 412.46 us, 0.43%
latency, 1.95 GFLOPS, )
          )
        (2): ResNetBottleNeckLayer(
          4.46 M, 18.98% Params, 1.75 GMACs, 5.34% MACs, 6.0 ms, 6.25% latency,
583.05 GFLOPS,
          (shortcut): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 18.6 us, 0.02%
latency, 0.0 FLOPS, )
          (layer): Sequential(
            4.46 M, 18.98% Params, 1.75 GMACs, 5.34% MACs, 5.39 ms, 5.62% latency,
648.74 GFLOPS,
              (0): ResNetConvLayer(
                1.05 M, 4.46% Params, 411.04 MMACs, 1.26% MACs, 1.62 ms, 1.68%
latency, 509.39 GFLOPS,
                (convolution): Conv2d(1.05 M, 4.46% Params, 411.04 MMACs, 1.26%
MACs, 145.67 us, 0.15% latency, 5.64 TFLOPS, 2048, 512, kernel_size=(1, 1),
stride=(1, 1), bias=False)
                (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 941.99 us, 0.98% latency, 426.13 MFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
                (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 422.24 us,
0.44% latency, 475.33 MFLOPS, )
              )
```

```
        (1): ResNetConvLayer(
          2.36 M, 10.04% Params, 924.84 MMACs, 2.83% MACs, 2.35 ms, 2.45%
latency, 786.69 GFLOPS,
          (convolution): Conv2d(2.36 M, 10.04% Params, 924.84 MMACs, 2.83%
MACs, 219.35 us, 0.23% latency, 8.43 TFLOPS, 512, 512, kernel_size=(3, 3),
stride=(1, 1), padding=(1, 1), bias=False)
          (normalization): BatchNorm2d(1.02 k, 0.00% Params, 0 MACs, 0.00%
MACs, 1.55 ms, 1.62% latency, 258.46 MFLOPS, 512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 442.27 us,
0.46% latency, 453.81 MFLOPS, )
        )
        (2): ResNetConvLayer(
          1.05 M, 4.48% Params, 411.04 MMACs, 1.26% MACs, 1.32 ms, 1.38%
latency, 621.7 GFLOPS,
          (convolution): Conv2d(1.05 M, 4.46% Params, 411.04 MMACs, 1.26%
MACs, 154.02 us, 0.16% latency, 5.34 TFLOPS, 512, 2048, kernel_size=(1, 1),
stride=(1, 1), bias=False)
          (normalization): BatchNorm2d(4.1 k, 0.02% Params, 0 MACs, 0.00%
MACs, 1.05 ms, 1.09% latency, 1.53 GFLOPS, 2048, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (activation): Identity(0, 0.00% Params, 0 MACs, 0.00% MACs, 22.65
us, 0.02% latency, 0.0 FLOPS, )
        )
      )
      (activation): ReLU(0, 0.00% Params, 0 MACs, 0.00% MACs, 430.35 us, 0.45%
latency, 1.87 GFLOPS, )
        )
      )
    )
  )
  (pooler): AdaptiveAvgPool2d(0, 0.00% Params, 0 MACs, 0.00% MACs, 321.63 us, 0.34%
latency, 2.5 GFLOPS, output_size=(1, 1))
)
--------------------------------------------------------------------------------
```