resnext101_32x8d architecture The layers highlighted in yellow are the ones whose weights are trained.

```
ResNet(
  (conv1): Conv2d(3, 64, kernel size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
  (relu): ReLU(inplace)
  (maxpool): MaxPool2d(kernel size=3, stride=2, padding=1, dilation=1, ceil mode=False)
  (layer1): Sequential(
    (0): Bottleneck(
      (conv1): Conv2d(64, 256, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-\overline{0}5, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(256, 256, kernel size=(3, 3), stride=(1, 1), padding=(1, 1), groups=32,
bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(256, 256, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
      (downsample): Sequential(
        (0): Conv2d(64, 256, kernel size=(1, 1), stride=(1, 1), bias=False)
        (1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
    (1): Bottleneck(
      (conv1): Conv2d(256, 256, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-0\overline{5}, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(256, 256, kernel size=(3, 3), stride=(1, 1), padding=(1, 1), groups=32,
bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(256, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (2): Bottleneck(
      (conv1): Conv2d(256, 256, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-0\overline{5}, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(256, 256, kernel size=(3, 3), stride=(1, 1), padding=(1, 1), groups=32,
bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(256, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
  (layer2): Sequential(
    (0): Bottleneck(
      (conv1): Conv2d(256, 512, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-0\overline{5}, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(512, 512, kernel size=(3, 3), stride=(2, 2), padding=(1, 1), groups=32,
bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(512, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
      (downsample): Sequential(
        (0): Conv2d(256, 512, kernel size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      )
    (1): Bottleneck(
      (conv1): Conv2d(512, 512, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(512, 512, kernel size=(3, 3), stride=(1, 1), padding=(1, 1), groups=32,
bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(512, 512, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
      (conv1): Conv2d(512, 512, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d (512, eps=1e-0\overline{5}, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(512, 512, kernel size=(3, 3), stride=(1, 1), padding=(1, 1), groups=32,
bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
```

```
(conv3): Conv2d(512, 512, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (3): Bottleneck(
      (\texttt{conv1}): \texttt{Conv2d}(\texttt{512}, \texttt{512}, \texttt{kernel\_size=(1, 1)}, \texttt{stride=(1, 1)}, \texttt{bias=False})
      (bn1): BatchNorm2d(512, eps=1e-0\overline{5}, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(512, 512, kernel size=(3, 3), stride=(1, 1), padding=(1, 1), groups=32,
bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(512, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
  (layer3): Sequential(
    (0): Bottleneck(
      (conv1): Conv2d(512, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(2, 2), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (downsample): Sequential(
        (0): Conv2d(512, 1024, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (1): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (2): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (3): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (4): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
    (5): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
```

```
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (6): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (7): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (\texttt{conv3}): \texttt{Conv2d}(1024, \ 1024, \ \texttt{kernel\_size=(1, 1)}, \ \texttt{stride=(1, 1)}, \ \texttt{bias=False})
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (8): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (9): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (10): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (11): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (12): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
```

```
(conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel\_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (14): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (15): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (16): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(1024, 1024, kernel\_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (17): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (18): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (relu): ReLU(inplace)
    (19): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(1024, 1024, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False) (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
    (20): Bottleneck(
      (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
      (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
      (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
```

```
(conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(1024, 1024, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
       (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
       (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=le-05, momentum=0.1, affine=True, track running stats=True)
       (relu): ReLU(inplace)
     (22): Bottleneck(
       (conv1): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
       (bn1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
       (conv2): Conv2d(1024, 1024, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
       (bn2): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
       (conv3): Conv2d(1024, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
       (relu): ReLU(inplace)
  (layer4): Sequential(
     (0): Bottleneck(
       (conv1): Conv2d(1024, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
       (conv2): Conv2d(2048, 2048, kernel size=(3, 3), stride=(2, 2), padding=(1, 1)
       (bn2): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
       (conv3): Conv2d(2048, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
       (relu): ReLU(inplace)
       (downsample): Sequential(
         (0): Conv2d(1024, 2048, kernel_size=(1, 1), stride=(2, 2), bias=False)
(1): BatchNorm2d(2048, eps=le-05, momentum=0.1, affine=True, track_running_stats=True)
     (1): Bottleneck(
       (conv1): Conv2d(2048, 2048, kernel size=(1, 1), stride=(1, 1), bias=False)
       (bn1): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
       (conv2): Conv2d(2048, 2048, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
       (bn2): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
       (conv3): Conv2d(2048, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
       (relu): ReLU(inplace)
     (2): Bottleneck (
       (conv1): Conv2d(2048, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
       (bn1): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
       (conv2): Conv2d(2048, 2048, kernel size=(3, 3), stride=(1, 1), padding=(1, 1),
groups=32, bias=False)
       (bn2): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
       (conv3): Conv2d(2048, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track running stats=True)
       (relu): ReLU(inplace)
  (avgpool): AdaptiveAvgPool2d(output_size=(1, 1))
  (fc): Linear(in_features=2048, out_features=1000, bias=True)
```