

Assignment 6 – Description of the (C)NN architectures

1 Own CNN

This is the model I designed for training on CIFAR10. It achieved over 97% accuracy score after running for 41 epochs on the setup I provided (which is shown in the proof documents mentioned in the README file).

1.1 Model's architecture

The following layers have been created for this model:

```
Model(  
    (layers): Sequential(  
      (0): Conv2d(3, 50, kernel_size=(3, 3), stride=(1, 1), padding=same  
      )  
      (1): ReLU()  
      (2): Conv2d(50, 64, kernel_size=(3, 3), stride=(1, 1), padding=  
      same)  
      (3): ReLU()  
      (4): MaxPool2d(kernel_size=(2, 2), stride=(2, 2), padding=0,  
      dilation=1, ceil_mode=False)  
      (5): Dropout(p=0.2, inplace=False)  
      (6): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=  
      same)  
      (7): ReLU()  
      (8): MaxPool2d(kernel_size=(2, 2), stride=(2, 2), padding=0,  
      dilation=1, ceil_mode=False)  
      (9): Dropout(p=0.2, inplace=False)  
      (10): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=  
      same)  
      (11): ReLU()  
      (12): MaxPool2d(kernel_size=(2, 2), stride=(2, 2), padding=0,  
      dilation=1, ceil_mode=False)
```

```

(13): Dropout(p=0.2, inplace=False)
(14): Flatten(start_dim=1, end_dim=-1)
(15): Linear(in_features=4096, out_features=500, bias=True)
(16): ReLU()
(17): Dropout(p=0.2, inplace=False)
(18): Linear(in_features=500, out_features=250, bias=True)
(19): ReLU()
(20): Dropout(p=0.2, inplace=False)
(21): Linear(in_features=250, out_features=10, bias=True)
(22): Softmax(dim=1)
)
(loss): CrossEntropyLoss()
)

```

The reason I have chained three blocks of type *Conv2d* – *ReLU* – *MaxPool2d* was the inspiration I got from VGG-16 well-known model, which I prevented from overfitting by dropout layers inserted amongst the convolutional ones (at the end of each such block). I have applied the same improvement with dropout to the last linear layers, whose purpose is to incorporate characteristics regarding the whole image and to globally connect the local information acquired from convolutions and max-pooling, through flow paths between the input neurons and the output classification.

1.2 Forward function

As discussed during the labs, the **Dropout** layer supports generalization and ensures 20% of (random) neurons on a particular layer are deactivated, by creating a mask for every input batch and setting the neurons to 0 before feeding them into the next layer. The **Conv2d** layers propagate a 3x3 filter over the image, replacing each pixel (since 1x1 stride is provided) with the weighted sum of the pixels amongst the 3x3 region surrounding it. The pixels whose surrounding region exceeds the representation's borders are discarded. The weights are dictated by the overlapping cells in the filter. The **MaxPool2d** layers extract the maximum from every 2x2 patch amongst the image, keeping a stride of 2x2 for the patch shifts. The use of pooling layers is to adress the issue of retaining insignificant details of the image, such as the position of specific objects; these details might reduce the power of generalization (even for basic augmentations, such as small rotations), which is why they are summarized by pooling layers into higher-level structural elements. Other

trials of mine also included **batch normalization** and **average pooling** layers, the latter aiming to extract the average of previously described patches.

The way I sequenced more convolutional layers allowed the model to define a pretty precise trajectory leading from low-level features (such as lines) to high-level features (such as shapes and specific objects).

1.3 Gradient flow

Taking into account the **convolutional** layers at first, backpropagation consists in adjusting both the input and the filter. Computing gradient is still based on the chain rule and relies on convolutions: the gradient with respect to the filter relies on the convolution between the input and the loss gradient with respect to the output; the gradient with respect to the input relies on the convolution between the flipped (180-degree rotated) filter and the loss gradient with respect to the output. To summarize, both the forward pass and the backpropagation step of a convolutional layer are convolutions. Since the pooling layers don't have any weights, we need to find the gradient of the error with respect to the input only. As far as **max pooling** layers are concerned, the gradient from the next layer is passed back to only that neuron which achieved the maximum value (and its derivative is the next to be computed further), whilst all the other neurons get zero gradient. This way, the gradients are "routed" towards the most productive features. If we switch to **average pooling** layers, the gradients are distributed back to the original input shape with respect to the number of elements which have been pooled.

2 PyramidNet

I used PyramidNet as a second variant of achieving over 97% accuracy score on CIFAR10. This model is based on rephrasing the principle of residual networks, which states that *the feature map dimension (i.e., the number of channels) is sharply increased at downsampling locations*, into a new working principle, which assumes that instead of sharply increasing the feature map dimension at units that perform downsampling, *the feature map dimension at all units are gradually increased so as to involve as many locations as possible*. The forward function and gradient flows for the various layers possess the same explanations as provided before for the first model, which is why only the model's architecture is described below.

2.1 Model's architecture

```
PyramidNet(  
    (loss): CrossEntropyLoss()  
    (conv1): Conv2d(3, 16, kernel_size=(3, 3), stride=(1, 1), padding  
        =(1, 1), bias=False)  
    (bn1): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,  
        track_running_stats=True)  
    (layer1): Sequential(  
        (0): Bottleneck(  
            (bn1): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,  
                track_running_stats=True)  
            (conv1): Conv2d(16, 16, kernel_size=(1, 1), stride=(1, 1), bias=  
                False)  
            (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,  
                track_running_stats=True)  
            (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),  
                padding=(1, 1), bias=False)  
            (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,  
                track_running_stats=True)  
            (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=  
                False)  
            (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,  
                track_running_stats=True)  
            (relu): ReLU(inplace=True)  
        )  
        (1): Bottleneck(  
            (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,  
                track_running_stats=True)  
            (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=  
                False)  
            (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,  
                track_running_stats=True)
```

```

(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(2): Bottleneck(
(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(3): Bottleneck(
(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
False)

```

```

(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(4): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (rel): ReLU(inplace=True)
)
(5): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)

```

```

(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(6): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(7): Bottleneck(

```

```

(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(relu): ReLU(inplace=True)
)
(8): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```



```

)
(9): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(10): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)

```

```

    (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(11): Bottleneck(
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(12): Bottleneck(
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)

```

```

(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
(rel): ReLU(inplace=True)
)
(13): Bottleneck(
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (rel): ReLU(inplace=True)
)
(14): Bottleneck(
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)

```

```

(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(15): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (rel): ReLU(inplace=True)
)
(16): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)

```

```

(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(17): Bottleneck(
(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(18): Bottleneck(
(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
False)

```

```

(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(19): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (rel): ReLU(inplace=True)
)
(20): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)

```

```

(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(21): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(22): Bottleneck(

```

```

(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(23): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (rel): ReLU(inplace=True)
)

```



```

)
(24): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(25): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)

```

```

        (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
            track_running_stats=True)
        (relu): ReLU(inplace=True)
    )
)
(layer2): Sequential(
  (0): Bottleneck(
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(2, 2),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
    (downsample): AvgPool2d(kernel_size=(2, 2), stride=(2, 2),
        padding=0)
  )
  (1): Bottleneck(
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)

```

```

(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(2): Bottleneck(
(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(3): Bottleneck(
(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
False)

```

```

(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(relu): ReLU(inplace=True)
)
(4): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(5): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)

```

```

(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(6): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(7): Bottleneck(

```

```

(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(8): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)

```

```

)
(9): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(10): Bottleneck(
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
    False)

```

```

    (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(11): Bottleneck(
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(16, 16, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(12): Bottleneck(
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(64, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)

```



```

(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
(rel): ReLU(inplace=True)
)
(13): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (rel): ReLU(inplace=True)
)
(14): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)

```

```

(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(15): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (rel): ReLU(inplace=True)
)
(16): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)

```

```

(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(17): Bottleneck(
(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(18): Bottleneck(
(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
False)

```

```

(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(19): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(20): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)

```

```

(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(21): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(22): Bottleneck(

```

```

(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(relu): ReLU(inplace=True)
)
(23): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

)
(24): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(25): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)

```

```

        (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
            track_running_stats=True)
        (relu): ReLU(inplace=True)
    )
)
(layer3): Sequential(
  (0): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(2, 2),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
    (downsample): AvgPool2d(kernel_size=(2, 2), stride=(2, 2),
        padding=0)
  )
  (1): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)

```



```

(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(2): Bottleneck(
(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(3): Bottleneck(
(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
False)

```

```

    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(4): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(5): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)

```

```

(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(6): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(7): Bottleneck(

```

```

(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(8): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)

```

```

)
(9): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(10): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)

```

```

    (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(11): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(12): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)

```

```

(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
(rel): ReLU(inplace=True)
)
(13): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)
    (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (rel): ReLU(inplace=True)
)
(14): Bottleneck(
    (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
        False)
    (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
        track_running_stats=True)
    (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
        padding=(1, 1), bias=False)

```

```

(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(15): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (rel): ReLU(inplace=True)
)
(16): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)

```



```

(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(17): Bottleneck(
(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(rel): ReLU(inplace=True)
)
(18): Bottleneck(
(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
False)

```

```

(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(19): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (rel): ReLU(inplace=True)
)
(20): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)

```

```

(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(rel): ReLU(inplace=True)
)
(21): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(22): Bottleneck(

```

```

(bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
  padding=(1, 1), bias=False)
(bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
  False)
(bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
  track_running_stats=True)
(relu): ReLU(inplace=True)
)
(23): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

)
(24): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(25): Bottleneck(
  (bn1): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv1): Conv2d(68, 17, kernel_size=(1, 1), stride=(1, 1), bias=
    False)
  (bn2): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv2): Conv2d(17, 17, kernel_size=(3, 3), stride=(1, 1),
    padding=(1, 1), bias=False)
  (bn3): BatchNorm2d(17, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
  (conv3): Conv2d(17, 68, kernel_size=(1, 1), stride=(1, 1), bias=
    False)

```

```

        (bn4): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
            track_running_stats=True)
        (relu): ReLU(inplace=True)
    )
)
(bn_final): BatchNorm2d(68, eps=1e-05, momentum=0.1, affine=True,
    track_running_stats=True)
(relu_final): ReLU(inplace=True)
(avgpool): AvgPool2d(kernel_size=8, stride=8, padding=0)
(fc): Linear(in_features=68, out_features=10, bias=True)
)

```

3 ResNet50 (pretrained)

The most notable concept regarding this type of network, which is the residual block, is based on the working principles of skip/shortcut connections.

3.1 Model's architecture

The ResNet with depth = 50 architecture is different from the architecture at <https://i.stack.imgur.com/XTo6Q.png> – which is ResNet34 – in that it replaces each 2-layer block in it with a 3-layer bottleneck block; this block consists in 1x1, 3x3 and 1x1 convolutions, where the 1x1 layers are responsible for reducing and then increasing dimensions, leaving the 3x3 layer with smaller dimensions.