

Model Architecture

```
DenseNet_Cifar(
  (features): Sequential(
    (conv0): Conv2d(3, 24, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    (denseblock1): _DenseBlock(
      (denselayer1): _DenseLayer(
        (norm1): BatchNorm2d(24, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(24, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      )
      (denselayer2): _DenseLayer(
        (norm1): BatchNorm2d(36, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(36, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      )
      (denselayer3): _DenseLayer(
        (norm1): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(48, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      )
      (denselayer4): _DenseLayer(
        (norm1): BatchNorm2d(60, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(60, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      )
      (denselayer5): _DenseLayer(
        (norm1): BatchNorm2d(72, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(72, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
```

```

(norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu2): ReLU(inplace=True)
(conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer6): _DenseLayer(
  (norm1): BatchNorm2d(84, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(84, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer7): _DenseLayer(
  (norm1): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(96, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer8): _DenseLayer(
  (norm1): BatchNorm2d(108, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(108, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer9): _DenseLayer(
  (norm1): BatchNorm2d(120, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(120, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer10): _DenseLayer(
  (norm1): BatchNorm2d(132, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(132, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)

```

```

    (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer11): _DenseLayer(
  (norm1): BatchNorm2d(144, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(144, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer12): _DenseLayer(
  (norm1): BatchNorm2d(156, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(156, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer13): _DenseLayer(
  (norm1): BatchNorm2d(168, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(168, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer14): _DenseLayer(
  (norm1): BatchNorm2d(180, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(180, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer15): _DenseLayer(
  (norm1): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(192, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)

```

```

(denselayer16): _DenseLayer(
  (norm1): BatchNorm2d(204, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(204, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
)
(transition1): _Transition(
  (norm): BatchNorm2d(216, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
  (conv): Conv2d(216, 108, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (pool): AvgPool2d(kernel_size=2, stride=2, padding=0)
)
(denseblock2): _DenseBlock(
  (denselayer1): _DenseLayer(
    (norm1): BatchNorm2d(108, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(108, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  )
  (denselayer2): _DenseLayer(
    (norm1): BatchNorm2d(120, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(120, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  )
  (denselayer3): _DenseLayer(
    (norm1): BatchNorm2d(132, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(132, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  )
  (denselayer4): _DenseLayer(
    (norm1): BatchNorm2d(144, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

```

```

(relu1): ReLU(inplace=True)
(conv1): Conv2d(144, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
(norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu2): ReLU(inplace=True)
(conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer5): _DenseLayer(
  (norm1): BatchNorm2d(156, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(156, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer6): _DenseLayer(
  (norm1): BatchNorm2d(168, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(168, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer7): _DenseLayer(
  (norm1): BatchNorm2d(180, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(180, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer8): _DenseLayer(
  (norm1): BatchNorm2d(192, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(192, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer9): _DenseLayer(
  (norm1): BatchNorm2d(204, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(204, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)

```

```
(norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(rel2): ReLU(inplace=True)
(conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer10): _DenseLayer(
  (norm1): BatchNorm2d(216, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(216, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer11): _DenseLayer(
  (norm1): BatchNorm2d(228, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(228, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer12): _DenseLayer(
  (norm1): BatchNorm2d(240, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(240, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer13): _DenseLayer(
  (norm1): BatchNorm2d(252, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(252, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer14): _DenseLayer(
  (norm1): BatchNorm2d(264, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(264, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
```

```

        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    )
    (denselayer15): _DenseLayer(
        (norm1): BatchNorm2d(276, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(276, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    )
    (denselayer16): _DenseLayer(
        (norm1): BatchNorm2d(288, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    )
)
(transition2): _Transition(
    (norm): BatchNorm2d(300, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv): Conv2d(300, 150, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (pool): AvgPool2d(kernel_size=2, stride=2, padding=0)
)
(denseblock3): _DenseBlock(
    (denselayer1): _DenseLayer(
        (norm1): BatchNorm2d(150, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(150, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    )
    (denselayer2): _DenseLayer(
        (norm1): BatchNorm2d(162, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(162, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    )
)

```

```

(denselayer3): _DenseLayer(
  (norm1): BatchNorm2d(174, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(174, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer4): _DenseLayer(
  (norm1): BatchNorm2d(186, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(186, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer5): _DenseLayer(
  (norm1): BatchNorm2d(198, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(198, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer6): _DenseLayer(
  (norm1): BatchNorm2d(210, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(210, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer7): _DenseLayer(
  (norm1): BatchNorm2d(222, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(222, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer8): _DenseLayer(
  (norm1): BatchNorm2d(234, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

```



```

(relu1): ReLU(inplace=True)
(conv1): Conv2d(234, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
(norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(relu2): ReLU(inplace=True)
(conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer9): _DenseLayer(
  (norm1): BatchNorm2d(246, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(246, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer10): _DenseLayer(
  (norm1): BatchNorm2d(258, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(258, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer11): _DenseLayer(
  (norm1): BatchNorm2d(270, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(270, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer12): _DenseLayer(
  (norm1): BatchNorm2d(282, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(282, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer13): _DenseLayer(
  (norm1): BatchNorm2d(294, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(294, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)

```

```

(norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(rel2): ReLU(inplace=True)
(conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer14): _DenseLayer(
  (norm1): BatchNorm2d(306, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(306, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer15): _DenseLayer(
  (norm1): BatchNorm2d(318, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(318, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
(denselayer16): _DenseLayer(
  (norm1): BatchNorm2d(330, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu1): ReLU(inplace=True)
  (conv1): Conv2d(330, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (norm2): BatchNorm2d(48, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu2): ReLU(inplace=True)
  (conv2): Conv2d(48, 12, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
)
)
(norm5): BatchNorm2d(342, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
(classifier): Linear(in_features=342, out_features=100, bias=True)
)

```

Layer (type)	Output Shape	Params	FLOPs(M+A) #
Conv2d-1	[1, 24, 32, 32]	648	1302528
BatchNorm2d-2	[1, 24, 32, 32]	96	49152
ReLU-3	[1, 24, 32, 32]	0	0
Conv2d-4	[1, 48, 32, 32]	1152	2310144
BatchNorm2d-5	[1, 48, 32, 32]	192	98304
ReLU-6	[1, 48, 32, 32]	0	0

Conv2d-7	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-8	[1, 36, 32, 32]	144	73728
ReLU-9	[1, 36, 32, 32]	0	0
Conv2d-10	[1, 48, 32, 32]	1728	3489792
BatchNorm2d-11	[1, 48, 32, 32]	192	98304
ReLU-12	[1, 48, 32, 32]	0	0
Conv2d-13	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-14	[1, 48, 32, 32]	192	98304
ReLU-15	[1, 48, 32, 32]	0	0
Conv2d-16	[1, 48, 32, 32]	2304	4669440
BatchNorm2d-17	[1, 48, 32, 32]	192	98304
ReLU-18	[1, 48, 32, 32]	0	0
Conv2d-19	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-20	[1, 60, 32, 32]	240	122880
ReLU-21	[1, 60, 32, 32]	0	0
Conv2d-22	[1, 48, 32, 32]	2880	5849088
BatchNorm2d-23	[1, 48, 32, 32]	192	98304
ReLU-24	[1, 48, 32, 32]	0	0
Conv2d-25	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-26	[1, 72, 32, 32]	288	147456
ReLU-27	[1, 72, 32, 32]	0	0
Conv2d-28	[1, 48, 32, 32]	3456	7028736
BatchNorm2d-29	[1, 48, 32, 32]	192	98304
ReLU-30	[1, 48, 32, 32]	0	0
Conv2d-31	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-32	[1, 84, 32, 32]	336	172032
ReLU-33	[1, 84, 32, 32]	0	0
Conv2d-34	[1, 48, 32, 32]	4032	8208384
BatchNorm2d-35	[1, 48, 32, 32]	192	98304
ReLU-36	[1, 48, 32, 32]	0	0
Conv2d-37	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-38	[1, 96, 32, 32]	384	196608
ReLU-39	[1, 96, 32, 32]	0	0
Conv2d-40	[1, 48, 32, 32]	4608	9388032
BatchNorm2d-41	[1, 48, 32, 32]	192	98304
ReLU-42	[1, 48, 32, 32]	0	0
Conv2d-43	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-44	[1, 108, 32, 32]	432	221184
ReLU-45	[1, 108, 32, 32]	0	0
Conv2d-46	[1, 48, 32, 32]	5184	10567680
BatchNorm2d-47	[1, 48, 32, 32]	192	98304
ReLU-48	[1, 48, 32, 32]	0	0

Conv2d-49	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-50	[1, 120, 32, 32]	480	245760
ReLU-51	[1, 120, 32, 32]	0	0
Conv2d-52	[1, 48, 32, 32]	5760	11747328
BatchNorm2d-53	[1, 48, 32, 32]	192	98304
ReLU-54	[1, 48, 32, 32]	0	0
Conv2d-55	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-56	[1, 132, 32, 32]	528	270336
ReLU-57	[1, 132, 32, 32]	0	0
Conv2d-58	[1, 48, 32, 32]	6336	12926976
BatchNorm2d-59	[1, 48, 32, 32]	192	98304
ReLU-60	[1, 48, 32, 32]	0	0
Conv2d-61	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-62	[1, 144, 32, 32]	576	294912
ReLU-63	[1, 144, 32, 32]	0	0
Conv2d-64	[1, 48, 32, 32]	6912	14106624
BatchNorm2d-65	[1, 48, 32, 32]	192	98304
ReLU-66	[1, 48, 32, 32]	0	0
Conv2d-67	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-68	[1, 156, 32, 32]	624	319488
ReLU-69	[1, 156, 32, 32]	0	0
Conv2d-70	[1, 48, 32, 32]	7488	15286272
BatchNorm2d-71	[1, 48, 32, 32]	192	98304
ReLU-72	[1, 48, 32, 32]	0	0
Conv2d-73	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-74	[1, 168, 32, 32]	672	344064
ReLU-75	[1, 168, 32, 32]	0	0
Conv2d-76	[1, 48, 32, 32]	8064	16465920
BatchNorm2d-77	[1, 48, 32, 32]	192	98304
ReLU-78	[1, 48, 32, 32]	0	0
Conv2d-79	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-80	[1, 180, 32, 32]	720	368640
ReLU-81	[1, 180, 32, 32]	0	0
Conv2d-82	[1, 48, 32, 32]	8640	17645568
BatchNorm2d-83	[1, 48, 32, 32]	192	98304
ReLU-84	[1, 48, 32, 32]	0	0
Conv2d-85	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-86	[1, 192, 32, 32]	768	393216
ReLU-87	[1, 192, 32, 32]	0	0
Conv2d-88	[1, 48, 32, 32]	9216	18825216
BatchNorm2d-89	[1, 48, 32, 32]	192	98304
ReLU-90	[1, 48, 32, 32]	0	0

Conv2d-91	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-92	[1, 204, 32, 32]	816	417792
ReLU-93	[1, 204, 32, 32]	0	0
Conv2d-94	[1, 48, 32, 32]	9792	20004864
BatchNorm2d-95	[1, 48, 32, 32]	192	98304
ReLU-96	[1, 48, 32, 32]	0	0
Conv2d-97	[1, 12, 32, 32]	5184	10604544
BatchNorm2d-98	[1, 216, 32, 32]	864	442368
ReLU-99	[1, 216, 32, 32]	0	0
Conv2d-100	[1, 108, 32, 32]	23328	47665152
AvgPool2d-101	[1, 108, 16, 16]	0	0
BatchNorm2d-102	[1, 108, 16, 16]	432	55296
ReLU-103	[1, 108, 16, 16]	0	0
Conv2d-104	[1, 48, 16, 16]	5184	2641920
BatchNorm2d-105	[1, 48, 16, 16]	192	24576
ReLU-106	[1, 48, 16, 16]	0	0
Conv2d-107	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-108	[1, 120, 16, 16]	480	61440
ReLU-109	[1, 120, 16, 16]	0	0
Conv2d-110	[1, 48, 16, 16]	5760	2936832
BatchNorm2d-111	[1, 48, 16, 16]	192	24576
ReLU-112	[1, 48, 16, 16]	0	0
Conv2d-113	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-114	[1, 132, 16, 16]	528	67584
ReLU-115	[1, 132, 16, 16]	0	0
Conv2d-116	[1, 48, 16, 16]	6336	3231744
BatchNorm2d-117	[1, 48, 16, 16]	192	24576
ReLU-118	[1, 48, 16, 16]	0	0
Conv2d-119	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-120	[1, 144, 16, 16]	576	73728
ReLU-121	[1, 144, 16, 16]	0	0
Conv2d-122	[1, 48, 16, 16]	6912	3526656
BatchNorm2d-123	[1, 48, 16, 16]	192	24576
ReLU-124	[1, 48, 16, 16]	0	0
Conv2d-125	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-126	[1, 156, 16, 16]	624	79872
ReLU-127	[1, 156, 16, 16]	0	0
Conv2d-128	[1, 48, 16, 16]	7488	3821568
BatchNorm2d-129	[1, 48, 16, 16]	192	24576
ReLU-130	[1, 48, 16, 16]	0	0
Conv2d-131	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-132	[1, 168, 16, 16]	672	86016

ReLU-133	[1, 168, 16, 16]	0	0
Conv2d-134	[1, 48, 16, 16]	8064	4116480
BatchNorm2d-135	[1, 48, 16, 16]	192	24576
ReLU-136	[1, 48, 16, 16]	0	0
Conv2d-137	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-138	[1, 180, 16, 16]	720	92160
ReLU-139	[1, 180, 16, 16]	0	0
Conv2d-140	[1, 48, 16, 16]	8640	4411392
BatchNorm2d-141	[1, 48, 16, 16]	192	24576
ReLU-142	[1, 48, 16, 16]	0	0
Conv2d-143	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-144	[1, 192, 16, 16]	768	98304
ReLU-145	[1, 192, 16, 16]	0	0
Conv2d-146	[1, 48, 16, 16]	9216	4706304
BatchNorm2d-147	[1, 48, 16, 16]	192	24576
ReLU-148	[1, 48, 16, 16]	0	0
Conv2d-149	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-150	[1, 204, 16, 16]	816	104448
ReLU-151	[1, 204, 16, 16]	0	0
Conv2d-152	[1, 48, 16, 16]	9792	5001216
BatchNorm2d-153	[1, 48, 16, 16]	192	24576
ReLU-154	[1, 48, 16, 16]	0	0
Conv2d-155	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-156	[1, 216, 16, 16]	864	110592
ReLU-157	[1, 216, 16, 16]	0	0
Conv2d-158	[1, 48, 16, 16]	10368	5296128
BatchNorm2d-159	[1, 48, 16, 16]	192	24576
ReLU-160	[1, 48, 16, 16]	0	0
Conv2d-161	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-162	[1, 228, 16, 16]	912	116736
ReLU-163	[1, 228, 16, 16]	0	0
Conv2d-164	[1, 48, 16, 16]	10944	5591040
BatchNorm2d-165	[1, 48, 16, 16]	192	24576
ReLU-166	[1, 48, 16, 16]	0	0
Conv2d-167	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-168	[1, 240, 16, 16]	960	122880
ReLU-169	[1, 240, 16, 16]	0	0
Conv2d-170	[1, 48, 16, 16]	11520	5885952
BatchNorm2d-171	[1, 48, 16, 16]	192	24576
ReLU-172	[1, 48, 16, 16]	0	0
Conv2d-173	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-174	[1, 252, 16, 16]	1008	129024

ReLU-175	[1, 252, 16, 16]	0	0
Conv2d-176	[1, 48, 16, 16]	12096	6180864
BatchNorm2d-177	[1, 48, 16, 16]	192	24576
ReLU-178	[1, 48, 16, 16]	0	0
Conv2d-179	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-180	[1, 264, 16, 16]	1056	135168
ReLU-181	[1, 264, 16, 16]	0	0
Conv2d-182	[1, 48, 16, 16]	12672	6475776
BatchNorm2d-183	[1, 48, 16, 16]	192	24576
ReLU-184	[1, 48, 16, 16]	0	0
Conv2d-185	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-186	[1, 276, 16, 16]	1104	141312
ReLU-187	[1, 276, 16, 16]	0	0
Conv2d-188	[1, 48, 16, 16]	13248	6770688
BatchNorm2d-189	[1, 48, 16, 16]	192	24576
ReLU-190	[1, 48, 16, 16]	0	0
Conv2d-191	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-192	[1, 288, 16, 16]	1152	147456
ReLU-193	[1, 288, 16, 16]	0	0
Conv2d-194	[1, 48, 16, 16]	13824	7065600
BatchNorm2d-195	[1, 48, 16, 16]	192	24576
ReLU-196	[1, 48, 16, 16]	0	0
Conv2d-197	[1, 12, 16, 16]	5184	2651136
BatchNorm2d-198	[1, 300, 16, 16]	1200	153600
ReLU-199	[1, 300, 16, 16]	0	0
Conv2d-200	[1, 150, 16, 16]	45000	23001600
AvgPool2d-201	[1, 150, 8, 8]	0	0
BatchNorm2d-202	[1, 150, 8, 8]	600	19200
ReLU-203	[1, 150, 8, 8]	0	0
Conv2d-204	[1, 48, 8, 8]	7200	918528
BatchNorm2d-205	[1, 48, 8, 8]	192	6144
ReLU-206	[1, 48, 8, 8]	0	0
Conv2d-207	[1, 12, 8, 8]	5184	662784
BatchNorm2d-208	[1, 162, 8, 8]	648	20736
ReLU-209	[1, 162, 8, 8]	0	0
Conv2d-210	[1, 48, 8, 8]	7776	992256
BatchNorm2d-211	[1, 48, 8, 8]	192	6144
ReLU-212	[1, 48, 8, 8]	0	0
Conv2d-213	[1, 12, 8, 8]	5184	662784
BatchNorm2d-214	[1, 174, 8, 8]	696	22272
ReLU-215	[1, 174, 8, 8]	0	0
Conv2d-216	[1, 48, 8, 8]	8352	1065984

BatchNorm2d-217	[1, 48, 8, 8]	192	6144
ReLU-218	[1, 48, 8, 8]	0	0
Conv2d-219	[1, 12, 8, 8]	5184	662784
BatchNorm2d-220	[1, 186, 8, 8]	744	23808
ReLU-221	[1, 186, 8, 8]	0	0
Conv2d-222	[1, 48, 8, 8]	8928	1139712
BatchNorm2d-223	[1, 48, 8, 8]	192	6144
ReLU-224	[1, 48, 8, 8]	0	0
Conv2d-225	[1, 12, 8, 8]	5184	662784
BatchNorm2d-226	[1, 198, 8, 8]	792	25344
ReLU-227	[1, 198, 8, 8]	0	0
Conv2d-228	[1, 48, 8, 8]	9504	1213440
BatchNorm2d-229	[1, 48, 8, 8]	192	6144
ReLU-230	[1, 48, 8, 8]	0	0
Conv2d-231	[1, 12, 8, 8]	5184	662784
BatchNorm2d-232	[1, 210, 8, 8]	840	26880
ReLU-233	[1, 210, 8, 8]	0	0
Conv2d-234	[1, 48, 8, 8]	10080	1287168
BatchNorm2d-235	[1, 48, 8, 8]	192	6144
ReLU-236	[1, 48, 8, 8]	0	0
Conv2d-237	[1, 12, 8, 8]	5184	662784
BatchNorm2d-238	[1, 222, 8, 8]	888	28416
ReLU-239	[1, 222, 8, 8]	0	0
Conv2d-240	[1, 48, 8, 8]	10656	1360896
BatchNorm2d-241	[1, 48, 8, 8]	192	6144
ReLU-242	[1, 48, 8, 8]	0	0
Conv2d-243	[1, 12, 8, 8]	5184	662784
BatchNorm2d-244	[1, 234, 8, 8]	936	29952
ReLU-245	[1, 234, 8, 8]	0	0
Conv2d-246	[1, 48, 8, 8]	11232	1434624
BatchNorm2d-247	[1, 48, 8, 8]	192	6144
ReLU-248	[1, 48, 8, 8]	0	0
Conv2d-249	[1, 12, 8, 8]	5184	662784
BatchNorm2d-250	[1, 246, 8, 8]	984	31488
ReLU-251	[1, 246, 8, 8]	0	0
Conv2d-252	[1, 48, 8, 8]	11808	1508352
BatchNorm2d-253	[1, 48, 8, 8]	192	6144
ReLU-254	[1, 48, 8, 8]	0	0
Conv2d-255	[1, 12, 8, 8]	5184	662784
BatchNorm2d-256	[1, 258, 8, 8]	1032	33024
ReLU-257	[1, 258, 8, 8]	0	0
Conv2d-258	[1, 48, 8, 8]	12384	1582080

BatchNorm2d-259	[1, 48, 8, 8]	192	6144
ReLU-260	[1, 48, 8, 8]	0	0
Conv2d-261	[1, 12, 8, 8]	5184	662784
BatchNorm2d-262	[1, 270, 8, 8]	1080	34560
ReLU-263	[1, 270, 8, 8]	0	0
Conv2d-264	[1, 48, 8, 8]	12960	1655808
BatchNorm2d-265	[1, 48, 8, 8]	192	6144
ReLU-266	[1, 48, 8, 8]	0	0
Conv2d-267	[1, 12, 8, 8]	5184	662784
BatchNorm2d-268	[1, 282, 8, 8]	1128	36096
ReLU-269	[1, 282, 8, 8]	0	0
Conv2d-270	[1, 48, 8, 8]	13536	1729536
BatchNorm2d-271	[1, 48, 8, 8]	192	6144
ReLU-272	[1, 48, 8, 8]	0	0
Conv2d-273	[1, 12, 8, 8]	5184	662784
BatchNorm2d-274	[1, 294, 8, 8]	1176	37632
ReLU-275	[1, 294, 8, 8]	0	0
Conv2d-276	[1, 48, 8, 8]	14112	1803264
BatchNorm2d-277	[1, 48, 8, 8]	192	6144
ReLU-278	[1, 48, 8, 8]	0	0
Conv2d-279	[1, 12, 8, 8]	5184	662784
BatchNorm2d-280	[1, 306, 8, 8]	1224	39168
ReLU-281	[1, 306, 8, 8]	0	0
Conv2d-282	[1, 48, 8, 8]	14688	1876992
BatchNorm2d-283	[1, 48, 8, 8]	192	6144
ReLU-284	[1, 48, 8, 8]	0	0
Conv2d-285	[1, 12, 8, 8]	5184	662784
BatchNorm2d-286	[1, 318, 8, 8]	1272	40704
ReLU-287	[1, 318, 8, 8]	0	0
Conv2d-288	[1, 48, 8, 8]	15264	1950720
BatchNorm2d-289	[1, 48, 8, 8]	192	6144
ReLU-290	[1, 48, 8, 8]	0	0
Conv2d-291	[1, 12, 8, 8]	5184	662784
BatchNorm2d-292	[1, 330, 8, 8]	1320	42240
ReLU-293	[1, 330, 8, 8]	0	0
Conv2d-294	[1, 48, 8, 8]	15840	2024448
BatchNorm2d-295	[1, 48, 8, 8]	192	6144
ReLU-296	[1, 48, 8, 8]	0	0
Conv2d-297	[1, 12, 8, 8]	5184	662784
BatchNorm2d-298	[1, 342, 8, 8]	1368	43776
Linear-299	[1, 100]	34300	68400
DenseNet_Cifar-300	[1, 100]	0	0

=====

Total parameters: 824,020 824.0K

Trainable parameters: 800,032

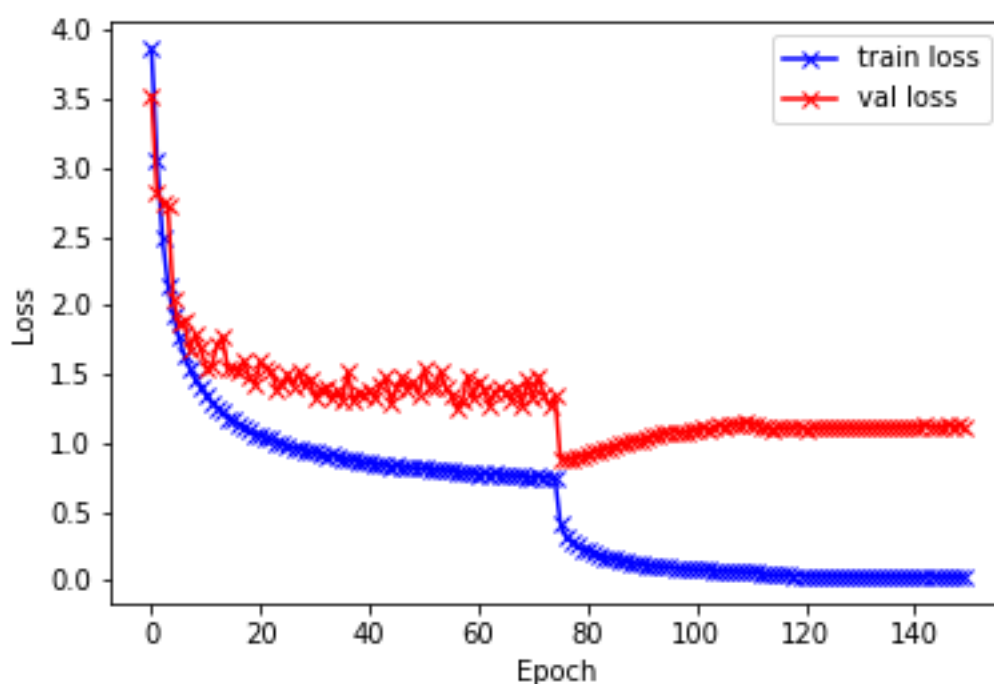
Non-trainable parameters: 23,988

Total flops(M) : 291,505,176 291.5M

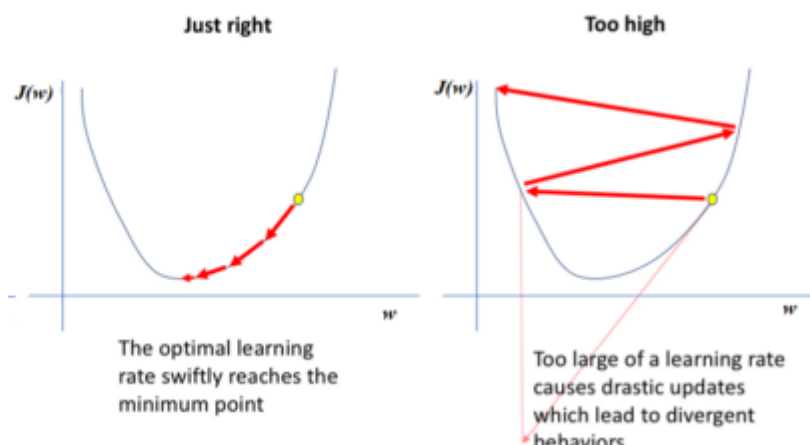
Total flops(M+A): 583,010,352 583.0M

Parameters size (MB): 3.14

Learning Curve



當 epoch 大約過 15 次之後，validation loss 與 training loss 越差越多，training loss 下降幅度較大，但 validation loss 卻沒什麼下降，這是 overfitting 的現象。當 epoch 達到 75 時，我將 learning rate 從原本的 0.1 縮小為 0.01，這讓 training loss 和 validation loss 都有大幅度下降。下方圖可以解釋，原本 learning rate 為 0.1，learning rate 太大讓 loss 無法收斂，而當 learning rate 減少一些後 step 比較小讓 loss 可以下降。



Method of Improve Accuracy

我原本是使用 Resnet 架構，但 accuracy 最多只能到 70% 左右，後來將架構改為 DenseNet 讓 accuracy 增加到 75%，其中使用 $\text{depth}=100$, $k=12$ ， k 是 DenseNet 中的參數，表示每個 layer 只會產生 k 個 feature maps，在 paper 中有紀錄其他 depth 和 k 用於 cifar100 的結果，若使用 $\text{depth}=250$, $k=24$ ，可以讓 accuracy 增加到 83% 左右，但因為深度太深造成 training 時間過長，使用 colab 會超出 GPU 使用限制，我有試著使用實驗室的 jupyter notebook，但常會發生斷線與 CUDA out of memory 的情況，就選擇使用 $\text{depth}=100$, $k=12$ 這組參數。

Data transform 的部分，在 train 階段使用了 RandomCrop/RandomHorizontalFlip/normalize，增加圖片的多樣性。在 training 時的 batch 實測過 64 與 128 發現 accuracy 沒有太大差別後便選擇 batch 為 64，而 paper 中使用的 epoch 為 300，我實測了 epoch 300/200/150 accuracy 也是差不多便選擇 epoch 為 150。cifar100 是要將圖片分類，選用最常用到的 cross entropy 做為 loss function。

Optimizer 則選用 SGD，learning rate 一開始設為 0.1，當 epoch 經過 50% 後變成 0.01，經過 75% 後變成 0.001，因此當 epoch 到 75 與 112 時，從 learning curve 可看出 loss 會產生變化。