DLCV hw1

Problem 1

1

Model structure of my Inception-v3 model:

```
ImgClassifier(
  (model): Inception3(
    (Conv2d_1a_3x3): BasicConv2d(
      (conv): Conv2d(3, 32, kernel_size=(3, 3), stride=(2, 2), bias=False)
      (bn): BatchNorm2d(32, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (Conv2d_2a_3x3): BasicConv2d(
      (conv): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), bias=False)
      (bn): BatchNorm2d(32, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (Conv2d_2b_3x3): BasicConv2d(
      (conv): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
      (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (maxpool1): MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1,
ceil_mode=False)
    (Conv2d_3b_1x1): BasicConv2d(
      (conv): Conv2d(64, 80, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn): BatchNorm2d(80, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (Conv2d_4a_3x3): BasicConv2d(
      (conv): Conv2d(80, 192, kernel_size=(3, 3), stride=(1, 1), bias=False)
      (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (maxpool2): MaxPool2d(kernel_size=3, stride=2, padding=0, dilation=1,
ceil_mode=False)
    (Mixed_5b): InceptionA(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(192, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch5x5_1): BasicConv2d(
        (conv): Conv2d(192, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(48, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch5x5_2): BasicConv2d(
```

```
        (conv): Conv2d(48, 64, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2),
bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_1): BasicConv2d(
        (conv): Conv2d(192, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_2): BasicConv2d(
        (conv): Conv2d(64, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (bn): BatchNorm2d(96, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_3): BasicConv2d(
        (conv): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (bn): BatchNorm2d(96, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(192, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(32, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (Mixed_5c): InceptionA(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch5x5_1): BasicConv2d(
        (conv): Conv2d(256, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(48, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch5x5_2): BasicConv2d(
        (conv): Conv2d(48, 64, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2),
bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_1): BasicConv2d(
        (conv): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_2): BasicConv2d(
        (conv): Conv2d(64, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (bn): BatchNorm2d(96, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_3): BasicConv2d(
```

```
        (conv): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (bn): BatchNorm2d(96, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (Mixed_5d): InceptionA(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(288, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch5x5_1): BasicConv2d(
        (conv): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(48, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch5x5_2): BasicConv2d(
        (conv): Conv2d(48, 64, kernel_size=(5, 5), stride=(1, 1), padding=(2, 2),
bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_1): BasicConv2d(
        (conv): Conv2d(288, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_2): BasicConv2d(
        (conv): Conv2d(64, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (bn): BatchNorm2d(96, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_3): BasicConv2d(
        (conv): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (bn): BatchNorm2d(96, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(288, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (Mixed_6a): InceptionB(
      (branch3x3): BasicConv2d(
        (conv): Conv2d(288, 384, kernel_size=(3, 3), stride=(2, 2), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_1): BasicConv2d(
```

```
        (conv): Conv2d(288, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(64, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_2): BasicConv2d(
        (conv): Conv2d(64, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (bn): BatchNorm2d(96, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_3): BasicConv2d(
        (conv): Conv2d(96, 96, kernel_size=(3, 3), stride=(2, 2), bias=False)
        (bn): BatchNorm2d(96, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (Mixed_6b): InceptionC(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_1): BasicConv2d(
        (conv): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(128, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_2): BasicConv2d(
        (conv): Conv2d(128, 128, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(128, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_3): BasicConv2d(
        (conv): Conv2d(128, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_1): BasicConv2d(
        (conv): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(128, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_2): BasicConv2d(
        (conv): Conv2d(128, 128, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(128, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_3): BasicConv2d(
        (conv): Conv2d(128, 128, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(128, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_4): BasicConv2d(
```

```
        (conv): Conv2d(128, 128, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(128, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_5): BasicConv2d(
        (conv): Conv2d(128, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (Mixed_6c): InceptionC(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_1): BasicConv2d(
        (conv): Conv2d(768, 160, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_2): BasicConv2d(
        (conv): Conv2d(160, 160, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_3): BasicConv2d(
        (conv): Conv2d(160, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_1): BasicConv2d(
        (conv): Conv2d(768, 160, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_2): BasicConv2d(
        (conv): Conv2d(160, 160, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_3): BasicConv2d(
        (conv): Conv2d(160, 160, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_4): BasicConv2d(
```

```
        (conv): Conv2d(160, 160, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_5): BasicConv2d(
        (conv): Conv2d(160, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (Mixed_6d): InceptionC(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_1): BasicConv2d(
        (conv): Conv2d(768, 160, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_2): BasicConv2d(
        (conv): Conv2d(160, 160, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_3): BasicConv2d(
        (conv): Conv2d(160, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_1): BasicConv2d(
        (conv): Conv2d(768, 160, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_2): BasicConv2d(
        (conv): Conv2d(160, 160, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_3): BasicConv2d(
        (conv): Conv2d(160, 160, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_4): BasicConv2d(
```

```
        (conv): Conv2d(160, 160, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(160, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_5): BasicConv2d(
        (conv): Conv2d(160, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (Mixed_6e): InceptionC(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_1): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_2): BasicConv2d(
        (conv): Conv2d(192, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7_3): BasicConv2d(
        (conv): Conv2d(192, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_1): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_2): BasicConv2d(
        (conv): Conv2d(192, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_3): BasicConv2d(
        (conv): Conv2d(192, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_4): BasicConv2d(
```

```
        (conv): Conv2d(192, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7dbl_5): BasicConv2d(
        (conv): Conv2d(192, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (AuxLogits): InceptionAux(
      (conv0): BasicConv2d(
        (conv): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(128, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (conv1): BasicConv2d(
        (conv): Conv2d(128, 768, kernel_size=(5, 5), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(768, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (fc): Linear(in_features=768, out_features=1000, bias=True)
    )
    (Mixed_7a): InceptionD(
      (branch3x3_1): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3_2): BasicConv2d(
        (conv): Conv2d(192, 320, kernel_size=(3, 3), stride=(2, 2), bias=False)
        (bn): BatchNorm2d(320, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7x3_1): BasicConv2d(
        (conv): Conv2d(768, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7x3_2): BasicConv2d(
        (conv): Conv2d(192, 192, kernel_size=(1, 7), stride=(1, 1), padding=(0,
3), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch7x7x3_3): BasicConv2d(
        (conv): Conv2d(192, 192, kernel_size=(7, 1), stride=(1, 1), padding=(3,
0), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
```

```
      (branch7x7x3_4): BasicConv2d(
        (conv): Conv2d(192, 192, kernel_size=(3, 3), stride=(2, 2), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (Mixed_7b): InceptionE(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(1280, 320, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(320, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3_1): BasicConv2d(
        (conv): Conv2d(1280, 384, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3_2a): BasicConv2d(
        (conv): Conv2d(384, 384, kernel_size=(1, 3), stride=(1, 1), padding=(0,
1), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3_2b): BasicConv2d(
        (conv): Conv2d(384, 384, kernel_size=(3, 1), stride=(1, 1), padding=(1,
0), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_1): BasicConv2d(
        (conv): Conv2d(1280, 448, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(448, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_2): BasicConv2d(
        (conv): Conv2d(448, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_3a): BasicConv2d(
        (conv): Conv2d(384, 384, kernel_size=(1, 3), stride=(1, 1), padding=(0,
1), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_3b): BasicConv2d(
        (conv): Conv2d(384, 384, kernel_size=(3, 1), stride=(1, 1), padding=(1,
0), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(1280, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
```

```
    (Mixed_7c): InceptionE(
      (branch1x1): BasicConv2d(
        (conv): Conv2d(2048, 320, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(320, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3_1): BasicConv2d(
        (conv): Conv2d(2048, 384, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3_2a): BasicConv2d(
        (conv): Conv2d(384, 384, kernel_size=(1, 3), stride=(1, 1), padding=(0,
1), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3_2b): BasicConv2d(
        (conv): Conv2d(384, 384, kernel_size=(3, 1), stride=(1, 1), padding=(1,
0), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_1): BasicConv2d(
        (conv): Conv2d(2048, 448, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(448, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_2): BasicConv2d(
        (conv): Conv2d(448, 384, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_3a): BasicConv2d(
        (conv): Conv2d(384, 384, kernel_size=(1, 3), stride=(1, 1), padding=(0,
1), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch3x3dbl_3b): BasicConv2d(
        (conv): Conv2d(384, 384, kernel_size=(3, 1), stride=(1, 1), padding=(1,
0), bias=False)
        (bn): BatchNorm2d(384, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
      (branch_pool): BasicConv2d(
        (conv): Conv2d(2048, 192, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn): BatchNorm2d(192, eps=0.001, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (avgpool): AdaptiveAvgPool2d(output_size=(1, 1))
    (dropout): Dropout(p=0.5, inplace=False)
    (fc): Sequential(
      (0): Linear(in_features=2048, out_features=1000, bias=True)
      (1): ReLU(inplace=True)
      (2): Dropout(p=0.5, inplace=False)
```

```
      (3): Linear(in_features=1000, out_features=50, bias=True)
    )
  )
)
```

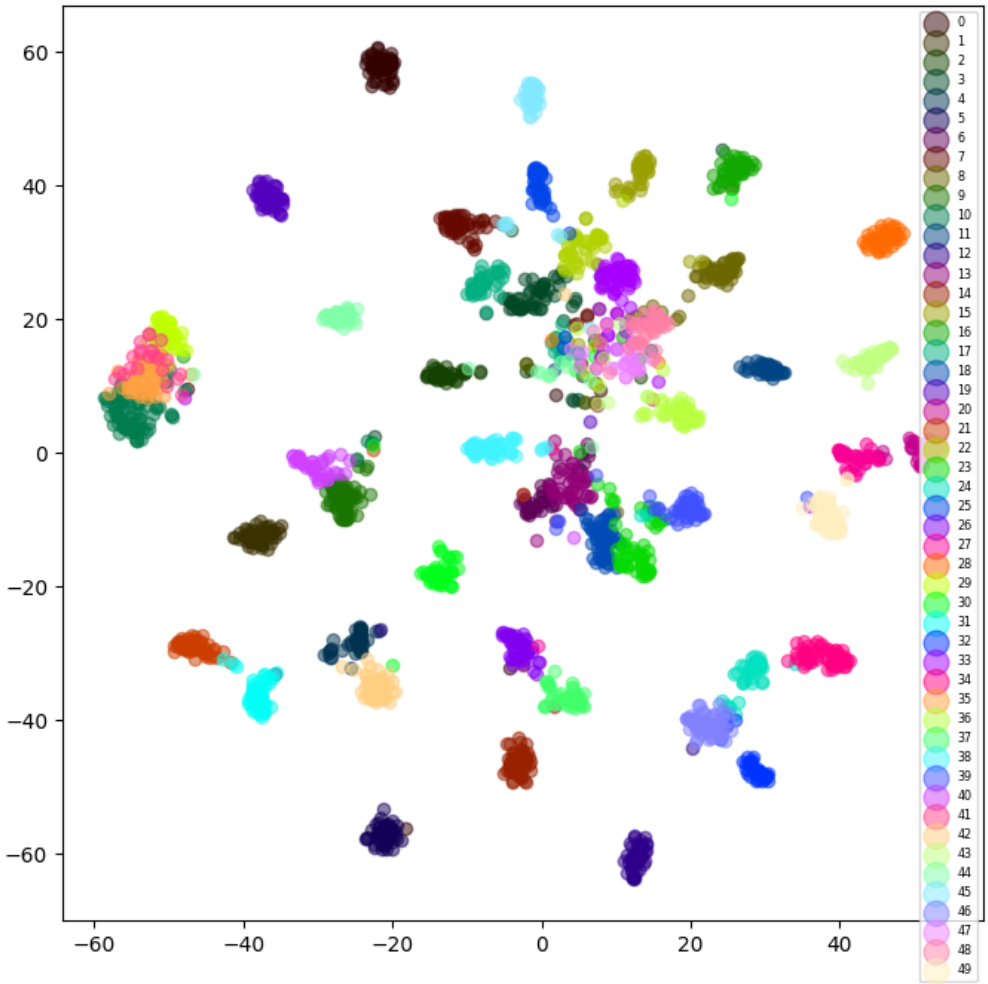
2

Accuracy: 0.8235999

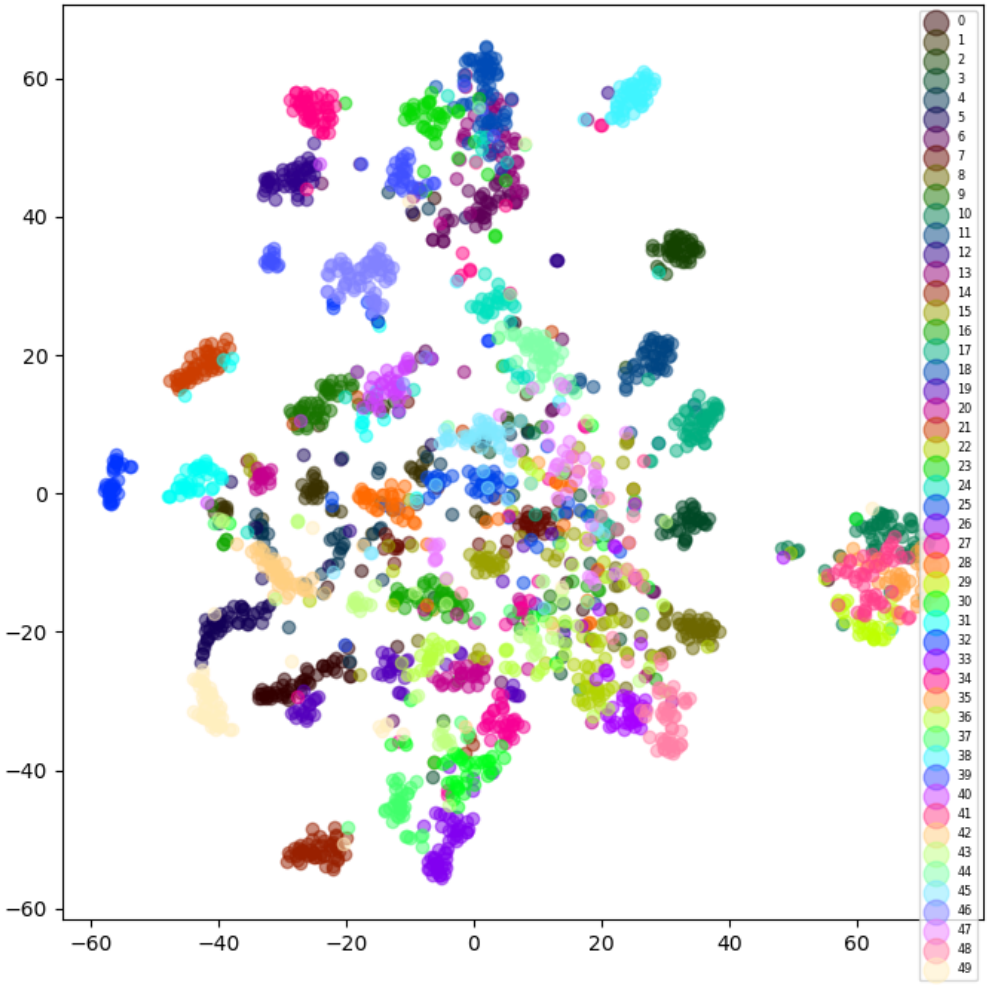
3

classifier的輸入為2048維的feature，對這些features representation做TSNE的結果如下，我嘗試了兩種作法：

(1) 直接用TSNE將2048維降到2維：



(2) 先用PCA將2048維降到10維，再用TSNE將10維降到2維



sklearn TSNE的document推薦先用PCA或TruncatedSVD降維到50維以內再做TSNE效果會比較好，說是可以把一些noise去除(同時當然也是為了加速)，但以hw1的例子看起來感覺直接做TSNE就好了。

Problem 2

1

My VGG16-FCN32's model structure:

```
VGG16_FCN32(
  (model): VGG(
    (features): Sequential(
      (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (1): ReLU(inplace=True)
      (2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (3): ReLU(inplace=True)
      (4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
      (5): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (6): ReLU(inplace=True)
      (7): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (8): ReLU(inplace=True)
```

```
      (9): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
      (10): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (11): ReLU(inplace=True)
      (12): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (13): ReLU(inplace=True)
      (14): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (15): ReLU(inplace=True)
      (16): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
      (17): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (18): ReLU(inplace=True)
      (19): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (20): ReLU(inplace=True)
      (21): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (22): ReLU(inplace=True)
      (23): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
      (24): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (25): ReLU(inplace=True)
      (26): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (27): ReLU(inplace=True)
      (28): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (29): ReLU(inplace=True)
      (30): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
    )
    (avgpool): AdaptiveAvgPool2d(output_size=(7, 7))
    (classifier): Sequential(
      (0): Linear(in_features=25088, out_features=4096, bias=True)
      (1): ReLU(inplace=True)
      (2): Dropout(p=0.5, inplace=False)
      (3): Linear(in_features=4096, out_features=4096, bias=True)
      (4): ReLU(inplace=True)
      (5): Dropout(p=0.5, inplace=False)
      (6): Linear(in_features=4096, out_features=1000, bias=True)
    )
  )
  (fc): Sequential(
    (0): Conv2d(512, 4096, kernel_size=(2, 2), stride=(1, 1))
    (1): ReLU(inplace=True)
    (2): Dropout2d(p=0.5, inplace=False)
    (3): Conv2d(4096, 4096, kernel_size=(1, 1), stride=(1, 1))
    (4): ReLU(inplace=True)
    (5): Dropout2d(p=0.5, inplace=False)
    (6): Conv2d(4096, 7, kernel_size=(1, 1), stride=(1, 1))
    (7): ConvTranspose2d(7, 7, kernel_size=(64, 64), stride=(32, 32), bias=False)
  )
)
```

2

Model structure of my improved model(VGG16-FCN16):

```
VGG16_FCN16(
  (vgg): VGG(
    (features): Sequential(
      (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (1): ReLU(inplace=True)
      (2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (3): ReLU(inplace=True)
      (4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
      (5): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (6): ReLU(inplace=True)
      (7): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (8): ReLU(inplace=True)
      (9): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
      (10): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (11): ReLU(inplace=True)
      (12): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (13): ReLU(inplace=True)
      (14): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (15): ReLU(inplace=True)
      (16): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
      (17): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (18): ReLU(inplace=True)
      (19): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (20): ReLU(inplace=True)
      (21): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (22): ReLU(inplace=True)
      (23): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
      (24): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (25): ReLU(inplace=True)
      (26): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (27): ReLU(inplace=True)
      (28): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
      (29): ReLU(inplace=True)
      (30): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
    )
    (avgpool): AdaptiveAvgPool2d(output_size=(7, 7))
    (classifier): Sequential(
      (0): Conv2d(512, 4096, kernel_size=(2, 2), stride=(1, 1))
      (1): ReLU(inplace=True)
      (2): Dropout2d(p=0.5, inplace=False)
      (3): Conv2d(4096, 4096, kernel_size=(1, 1), stride=(1, 1))
      (4): ReLU(inplace=True)
      (5): Dropout2d(p=0.5, inplace=False)
      (6): Conv2d(4096, 7, kernel_size=(1, 1), stride=(1, 1))
      (7): ConvTranspose2d(7, 512, kernel_size=(4, 4), stride=(2, 2), bias=False)
    )
  )
  (start_pool4): Sequential(
```

```
    (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (3): ReLU(inplace=True)
    (4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
    (5): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (6): ReLU(inplace=True)
    (7): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (8): ReLU(inplace=True)
    (9): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
    (10): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (11): ReLU(inplace=True)
    (12): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (13): ReLU(inplace=True)
    (14): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (15): ReLU(inplace=True)
    (16): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
    (17): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (18): ReLU(inplace=True)
    (19): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (20): ReLU(inplace=True)
    (21): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (22): ReLU(inplace=True)
    (23): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  )
  (pool4_pool5): Sequential(
    (0): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (3): ReLU(inplace=True)
    (4): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (5): ReLU(inplace=True)
    (6): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  )
  (upsample16): ConvTranspose2d(512, 7, kernel_size=(16, 16), stride=(16, 16),
bias=False)
)
```

3

Baseline performance of my VGG16-FCN32 model:
class #0 : 0.72276
class #1 : 0.87115
class #2 : 0.28123
class #3 : 0.78886
class #4 : 0.73544
class #5 : 0.63814
class #6 : 0.00000
Vailidation MIOU: 0.672928829305534

Performance of my improved model (VGG16-FCN16):
class #0 : 0.74053
class #1 : 0.87938
class #2 : 0.30541
class #3 : 0.80835
class #4 : 0.73479
class #5 : 0.65181
class #6 : 0.00000
Vailidation MIOU: 0.6867122977021437

4

拿improved model去inference的結果：