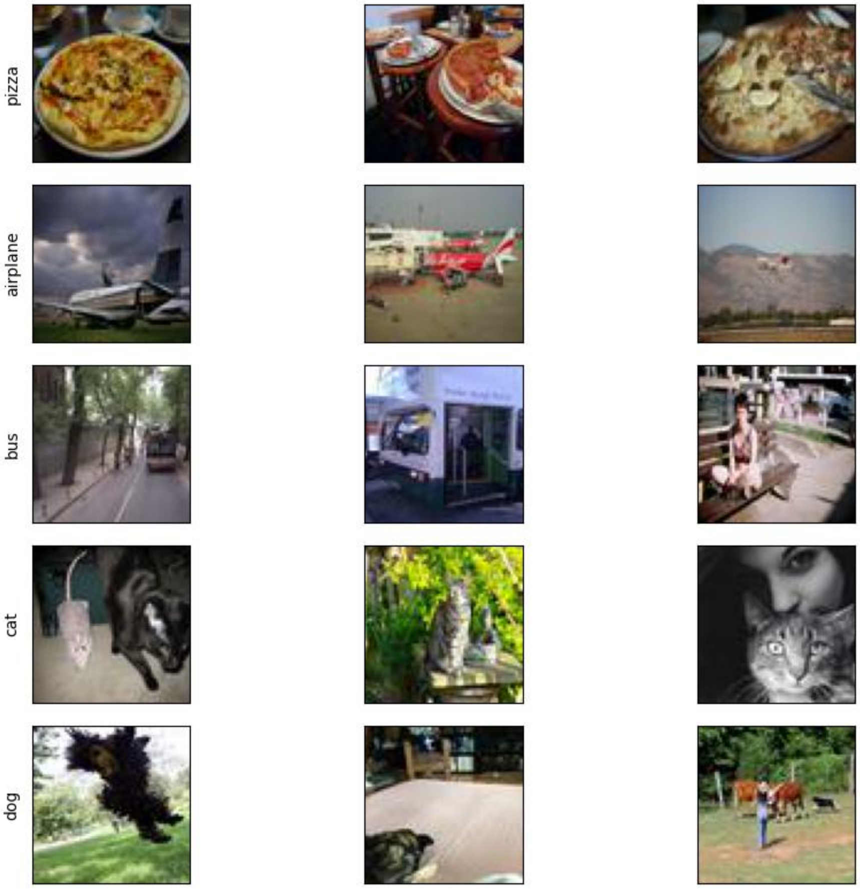
HW4 Report

Yuxin Sun

3.1 Creating Your Own Image Classification Dataset

In my DataSet, I sample 2000 unique images for each class. The training set and testing set are stored together in one DataSet, and I use `SubsetRandomSampler` in `torch.utils.data.sampler` to make sure there is no overlap between two sets.

Figure 1: resized image from COCO



3.2 Image Classification using CNNs – Training and Validation

parameters

In the following code, xxxx = 32*14*14:

Our input image is 64×64, and it changes as $64 \rightarrow 62 \rightarrow 31 \rightarrow 29 \rightarrow 14$, so the final image should have size 14×14. The final `Conv2d` has 32 channel, so xxxx=32*14*14.

xx = 5 because it is the number of classes we have.

```
class HW4Net1(nn.Module):
    def __init__(self):
        super(HW4Net1, self).__init__()
        self.conv1 = nn.Conv2d(3, 16, 3)
        self.pool = nn.MaxPool2d(2,2)
        self.conv2 = nn.Conv2d(16, 32, 3)
        self.fc1 = nn.Linear(xxxx, 64)
        self.fc2 = nn.Linear(64, xx)
```

Confusion matrix

Confusion matrices for 3 nets:

```
tensor([[   0.,   37.,  175.,  216., 1072.],
        [   0., 1094.,  236.,   44.,  126.],
        [   0.,  176., 1054.,   82.,  188.],
        [   0.,  110.,  243.,  563.,  584.],
        [   0.,  168.,  376.,  341.,  615.]])

tensor([[   0.,   12.,   85.,  133., 1270.],
        [   0.,  902.,  264.,   64.,  270.],
        [   0.,  100.,  963.,   70.,  367.],
        [   0.,   55.,  154.,  492.,  799.],
        [   0.,  102.,  242.,  275.,  881.]])

tensor([[   0.,   24.,   75.,  158., 1243.],
        [   0., 1100.,  184.,   61.,  155.],
        [   0.,  184.,  832.,  141.,  343.],
        [   0.,  177.,  178.,  434.,  711.],
        [   0.,  221.,  276.,  331.,  672.]])
```

Normalized Confusion matrices plot

Figure 2:

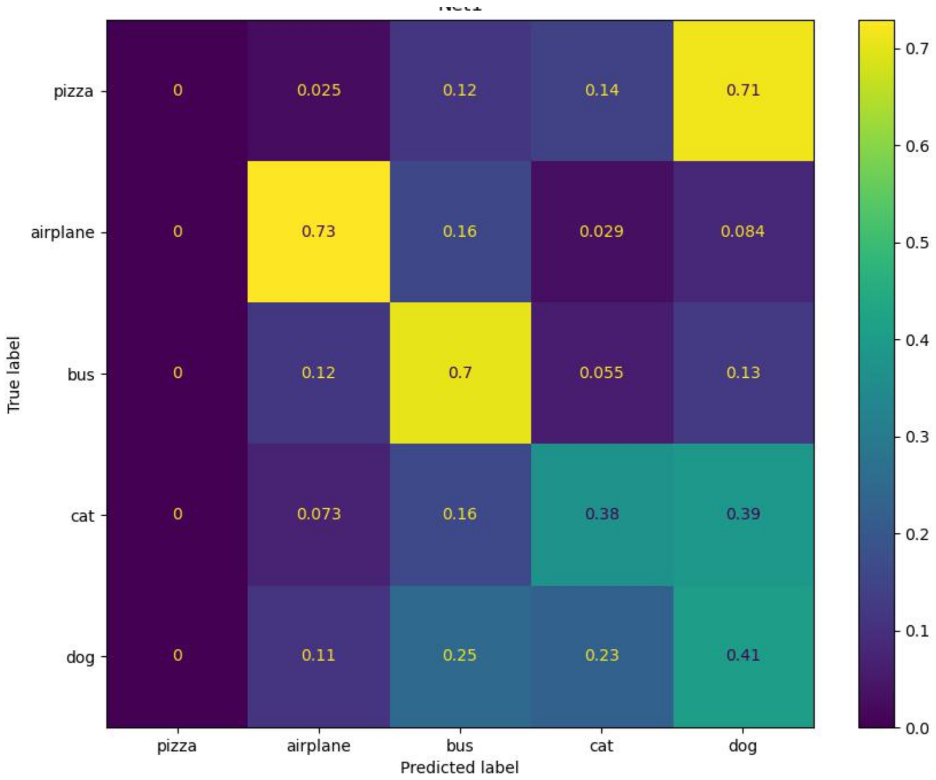


Figure 3:

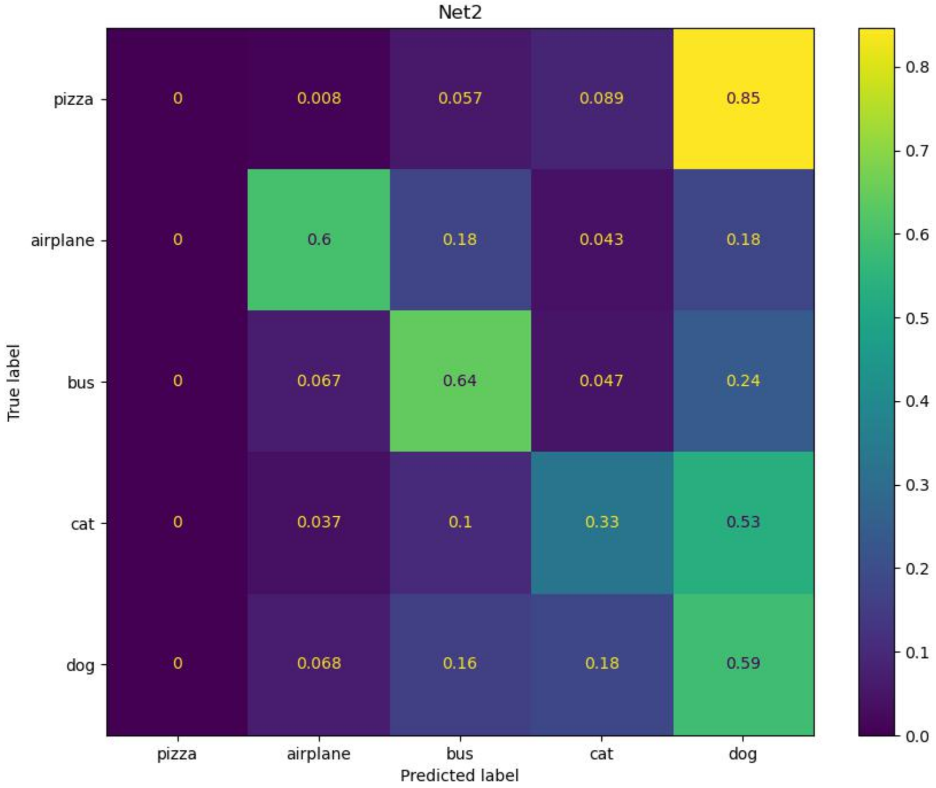


Figure 4:

Figure 4:

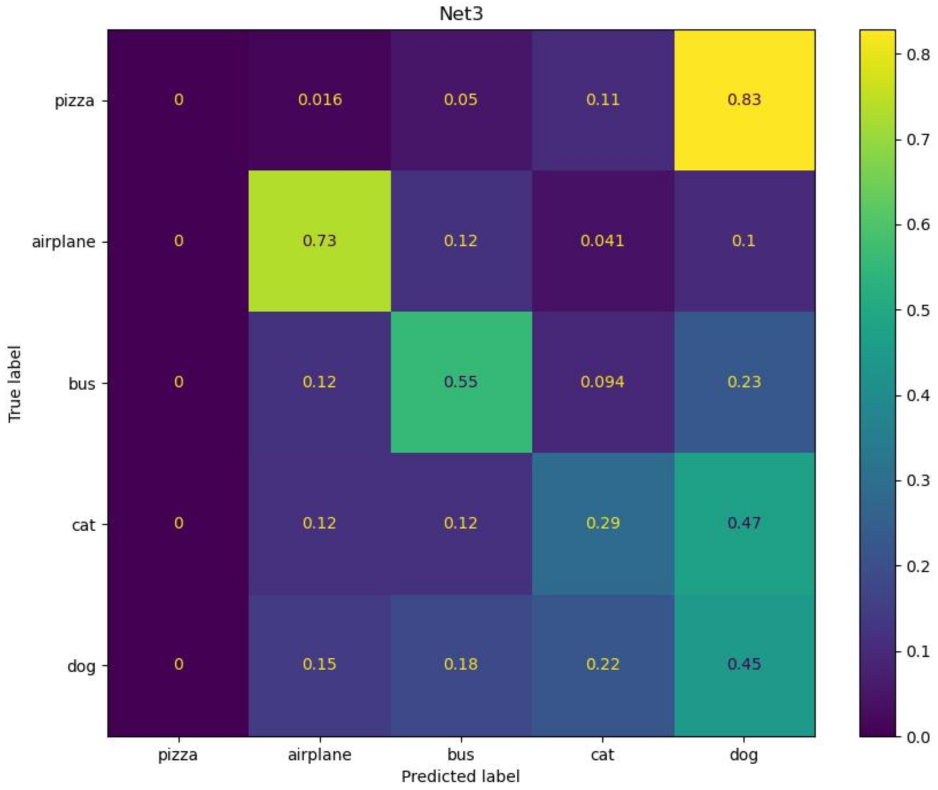
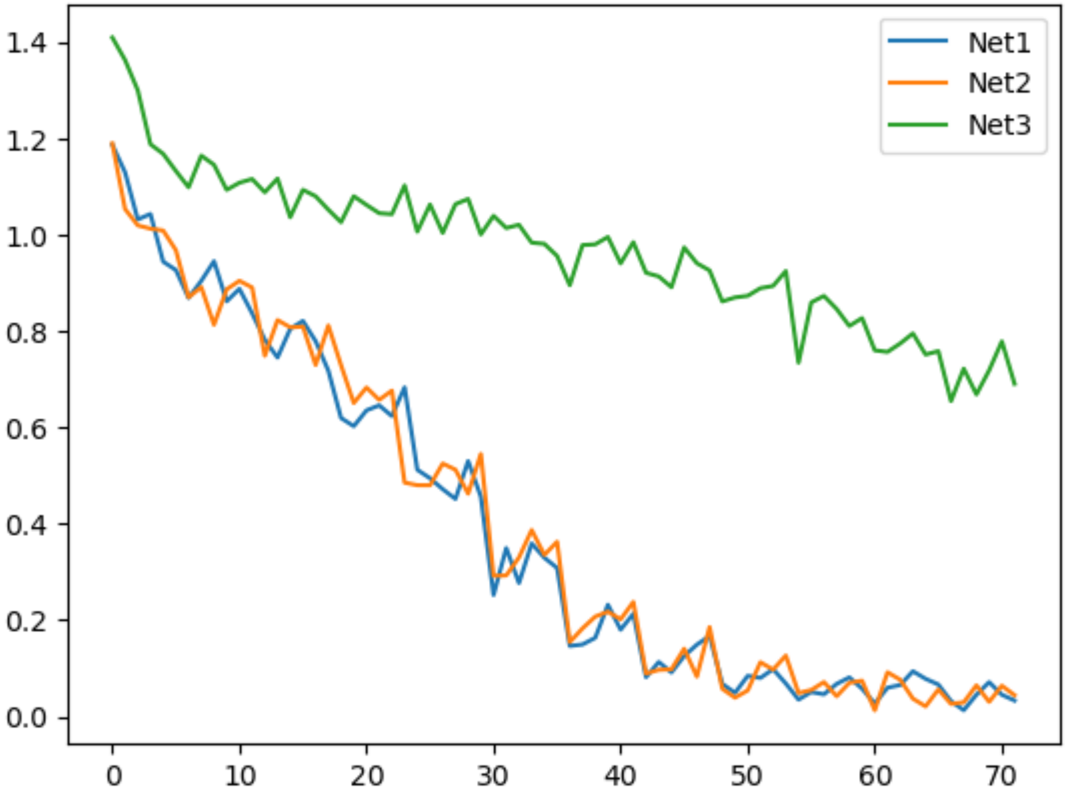


Figure 5: Training Loss vs #Batch/100



Questions

1. Does adding padding to the convolutional layers make a difference in classification performance?

A: Yes, the confusion matrix looks a little bit different. But adding padding doesn't make it performs better.

2. As you may have known, naively chaining a large number of layers can result in difficulties in training. This phenomenon is often referred to as vanishing gradient. Do you observe something like that in Net3?

A: I observed that the training loss of Net3 decreases slowly, compared to previous nets.

3. Compare the classification results by all three networks, which CNN do you think is the best performer?

A: I think Net 1 performs better, because its diagonal elements of confusion matrix are larger.

4. By observing your confusion matrices, which class or classes do you think are more difficult to correctly differentiate and why?

A: I think it is "pizza". I am not sure if I make some mistakes in my Net, because my Nets predict most "pizza" images as "dog".

5. What is one thing that you propose to make the classification performance better?

A: Maybe we can use Skip Connections method to avoid the possible "vanishing gradient".