

CNN-181

March 23, 2024

Acknowledgement:

- Code skeleton for this project is based on homework 4.
- For this project I try to train dataset on a new CNN model with Adamax optimizer.

```
[1]: %matplotlib inline
import matplotlib.pyplot as plt
import numpy as np
import torch
import torchvision
import torchvision.transforms as transforms
import torch.nn as nn
import torch.nn.functional as F
import torch.optim as optim
```

Prepare for Dataset

```
[2]: transform = transforms.Compose(
    [transforms.ToTensor(),
     transforms.Normalize((0.5, 0.5, 0.5), (0.5, 0.5, 0.5))])

trainset = torchvision.datasets.CIFAR10(root='./data', train=True,
                                         download=True, transform=transform)
trainloader = torch.utils.data.DataLoader(trainset, batch_size=4,
                                           shuffle=True, num_workers=2)

testset = torchvision.datasets.CIFAR10(root='./data', train=False,
                                         download=True, transform=transform)
testloader = torch.utils.data.DataLoader(testset, batch_size=4,
                                          shuffle=False, num_workers=2)

classes = ('plane', 'car', 'bird', 'cat',
           'deer', 'dog', 'frog', 'horse', 'ship', 'truck')
```

Files already downloaded and verified

Files already downloaded and verified

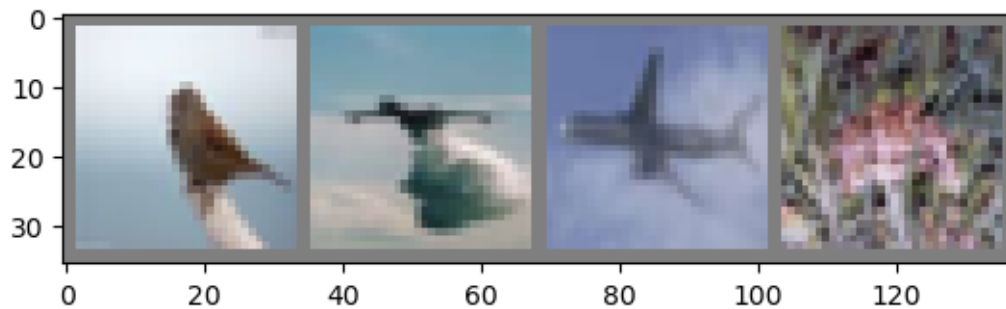
```
[3]: # The function to show an image.
def imshow(img):
    img = img / 2 + 0.5     # Unnormalize.
```

```

    npimg = img.numpy()
    plt.imshow(np.transpose(npimg, (1, 2, 0)))
    plt.show()

# Get some random training images.
dataiter = iter(trainloader)
images, labels = next(dataiter)
# Show images.
imshow(torchvision.utils.make_grid(images))
# Print labels.
print(' '.join('%5s' % classes[labels[j]] for j in range(4)))

```



bird plane plane frog

Choose a Device

```

[4]: # If there are GPUs, choose the first one for computing. Otherwise use CPU.
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
print(device)
# If 'cuda:0' is printed, it means GPU is available.

```

cpu

Network Definition

```

[5]: class Net(nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.conv1 = nn.Conv2d(3, 64, kernel_size=3, padding=1)
        self.relu1 = nn.ReLU()
        self.pool1 = nn.MaxPool2d(kernel_size=2, stride=2)

        self.conv2 = nn.Conv2d(64, 128, kernel_size=3, padding=1)
        self.relu2 = nn.ReLU()
        self.pool2 = nn.MaxPool2d(kernel_size=2, stride=2)

        self.conv3 = nn.Conv2d(128, 256, kernel_size=3, padding=1)

```

```

self.relu3 = nn.ReLU()
self.pool3 = nn.MaxPool2d(kernel_size=2, stride=2)

self.fc1 = nn.Linear(256 * 4 * 4, 1024)
self.relu4 = nn.ReLU()
self.fc2 = nn.Linear(1024, 10)

def forward(self, x):
    x = self.pool1(self.relu1(self.conv1(x)))
    x = self.pool2(self.relu2(self.conv2(x)))
    x = self.pool3(self.relu3(self.conv3(x)))
    x = x.view(-1, 256 * 4 * 4)
    x = self.relu4(self.fc1(x))
    x = self.fc2(x)
    return x

net = Net()
net.to(device)

```

```

[5]: Net(
  (conv1): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (relu1): ReLU()
  (pool1): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (conv2): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (relu2): ReLU()
  (pool2): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (conv3): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (relu3): ReLU()
  (pool3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (fc1): Linear(in_features=4096, out_features=1024, bias=True)
  (relu4): ReLU()
  (fc2): Linear(in_features=1024, out_features=10, bias=True)
)

```

Optimizer and Loss Function

```

[6]: # We use cross-entropy as loss function.
loss_func = nn.CrossEntropyLoss()
# We use stochastic gradient descent (SGD) as optimizer.
opt = optim.SGD(net.parameters(), lr=0.001, momentum=0.9)

```

Training Procedure

```

[10]: ## code is
avg_losses = [] # Avg. losses.
epochs = 10 # Total epochs.

```

```

print_freq = 100  # Print frequency.

for epoch in range(epochs):  # Loop over the dataset multiple times.
    running_loss = 0.0        # Initialize running loss.
    for i, data in enumerate(trainloader, 0):
        # Get the inputs.
        inputs, labels = data

        # Move the inputs to the specified device.
        inputs, labels = inputs.to(device), labels.to(device)

        # Zero the parameter gradients.
        opt.zero_grad()

        # Forward step.
        outputs = net(inputs)
        loss = loss_func(outputs, labels)

        # Backward step.
        loss.backward()

        # Optimization step (update the parameters).
        opt.step()

        # Print statistics.
        running_loss += loss.item()
        if i % print_freq == print_freq - 1: # Print every several mini-batches.
            avg_loss = running_loss / print_freq
            print('[epoch: {}, i: {:5d}] avg mini-batch loss: {:.3f}'.format(
                epoch, i, avg_loss))
            avg_losses.append(avg_loss)
            running_loss = 0.0

print('Finished Training.')

```

```

[epoch: 0, i:    99] avg mini-batch loss: 1.313
[epoch: 0, i:   199] avg mini-batch loss: 1.347
[epoch: 0, i:   299] avg mini-batch loss: 1.262
[epoch: 0, i:   399] avg mini-batch loss: 1.345
[epoch: 0, i:   499] avg mini-batch loss: 1.243
[epoch: 0, i:   599] avg mini-batch loss: 1.342
[epoch: 0, i:   699] avg mini-batch loss: 1.366
[epoch: 0, i:   799] avg mini-batch loss: 1.297
[epoch: 0, i:   899] avg mini-batch loss: 1.354
[epoch: 0, i:   999] avg mini-batch loss: 1.343
[epoch: 0, i:  1099] avg mini-batch loss: 1.352
[epoch: 0, i:  1199] avg mini-batch loss: 1.314
[epoch: 0, i:  1299] avg mini-batch loss: 1.251

```

[epoch: 0, i: 1399] avg mini-batch loss: 1.245
[epoch: 0, i: 1499] avg mini-batch loss: 1.340
[epoch: 0, i: 1599] avg mini-batch loss: 1.279
[epoch: 0, i: 1699] avg mini-batch loss: 1.184
[epoch: 0, i: 1799] avg mini-batch loss: 1.298
[epoch: 0, i: 1899] avg mini-batch loss: 1.335
[epoch: 0, i: 1999] avg mini-batch loss: 1.218
[epoch: 0, i: 2099] avg mini-batch loss: 1.255
[epoch: 0, i: 2199] avg mini-batch loss: 1.245
[epoch: 0, i: 2299] avg mini-batch loss: 1.220
[epoch: 0, i: 2399] avg mini-batch loss: 1.225
[epoch: 0, i: 2499] avg mini-batch loss: 1.243
[epoch: 0, i: 2599] avg mini-batch loss: 1.219
[epoch: 0, i: 2699] avg mini-batch loss: 1.253
[epoch: 0, i: 2799] avg mini-batch loss: 1.254
[epoch: 0, i: 2899] avg mini-batch loss: 1.170
[epoch: 0, i: 2999] avg mini-batch loss: 1.348
[epoch: 0, i: 3099] avg mini-batch loss: 1.198
[epoch: 0, i: 3199] avg mini-batch loss: 1.172
[epoch: 0, i: 3299] avg mini-batch loss: 1.190
[epoch: 0, i: 3399] avg mini-batch loss: 1.200
[epoch: 0, i: 3499] avg mini-batch loss: 1.121
[epoch: 0, i: 3599] avg mini-batch loss: 1.212
[epoch: 0, i: 3699] avg mini-batch loss: 1.069
[epoch: 0, i: 3799] avg mini-batch loss: 1.163
[epoch: 0, i: 3899] avg mini-batch loss: 1.264
[epoch: 0, i: 3999] avg mini-batch loss: 1.237
[epoch: 0, i: 4099] avg mini-batch loss: 1.160
[epoch: 0, i: 4199] avg mini-batch loss: 1.151
[epoch: 0, i: 4299] avg mini-batch loss: 1.153
[epoch: 0, i: 4399] avg mini-batch loss: 1.104
[epoch: 0, i: 4499] avg mini-batch loss: 1.204
[epoch: 0, i: 4599] avg mini-batch loss: 1.143
[epoch: 0, i: 4699] avg mini-batch loss: 1.178
[epoch: 0, i: 4799] avg mini-batch loss: 1.159
[epoch: 0, i: 4899] avg mini-batch loss: 1.170
[epoch: 0, i: 4999] avg mini-batch loss: 1.133
[epoch: 0, i: 5099] avg mini-batch loss: 1.183
[epoch: 0, i: 5199] avg mini-batch loss: 1.147
[epoch: 0, i: 5299] avg mini-batch loss: 1.116
[epoch: 0, i: 5399] avg mini-batch loss: 1.205
[epoch: 0, i: 5499] avg mini-batch loss: 1.105
[epoch: 0, i: 5599] avg mini-batch loss: 1.162
[epoch: 0, i: 5699] avg mini-batch loss: 1.085
[epoch: 0, i: 5799] avg mini-batch loss: 1.073
[epoch: 0, i: 5899] avg mini-batch loss: 1.163
[epoch: 0, i: 5999] avg mini-batch loss: 1.088
[epoch: 0, i: 6099] avg mini-batch loss: 1.082

[epoch: 0, i: 6199] avg mini-batch loss: 1.135
[epoch: 0, i: 6299] avg mini-batch loss: 1.033
[epoch: 0, i: 6399] avg mini-batch loss: 1.071
[epoch: 0, i: 6499] avg mini-batch loss: 1.142
[epoch: 0, i: 6599] avg mini-batch loss: 1.090
[epoch: 0, i: 6699] avg mini-batch loss: 1.084
[epoch: 0, i: 6799] avg mini-batch loss: 1.034
[epoch: 0, i: 6899] avg mini-batch loss: 1.053
[epoch: 0, i: 6999] avg mini-batch loss: 1.062
[epoch: 0, i: 7099] avg mini-batch loss: 1.144
[epoch: 0, i: 7199] avg mini-batch loss: 1.143
[epoch: 0, i: 7299] avg mini-batch loss: 1.119
[epoch: 0, i: 7399] avg mini-batch loss: 1.019
[epoch: 0, i: 7499] avg mini-batch loss: 1.131
[epoch: 0, i: 7599] avg mini-batch loss: 1.093
[epoch: 0, i: 7699] avg mini-batch loss: 1.079
[epoch: 0, i: 7799] avg mini-batch loss: 1.090
[epoch: 0, i: 7899] avg mini-batch loss: 1.039
[epoch: 0, i: 7999] avg mini-batch loss: 1.115
[epoch: 0, i: 8099] avg mini-batch loss: 1.015
[epoch: 0, i: 8199] avg mini-batch loss: 0.968
[epoch: 0, i: 8299] avg mini-batch loss: 0.976
[epoch: 0, i: 8399] avg mini-batch loss: 1.039
[epoch: 0, i: 8499] avg mini-batch loss: 1.069
[epoch: 0, i: 8599] avg mini-batch loss: 1.022
[epoch: 0, i: 8699] avg mini-batch loss: 1.079
[epoch: 0, i: 8799] avg mini-batch loss: 1.018
[epoch: 0, i: 8899] avg mini-batch loss: 1.036
[epoch: 0, i: 8999] avg mini-batch loss: 1.143
[epoch: 0, i: 9099] avg mini-batch loss: 1.022
[epoch: 0, i: 9199] avg mini-batch loss: 1.066
[epoch: 0, i: 9299] avg mini-batch loss: 0.918
[epoch: 0, i: 9399] avg mini-batch loss: 0.968
[epoch: 0, i: 9499] avg mini-batch loss: 1.043
[epoch: 0, i: 9599] avg mini-batch loss: 1.110
[epoch: 0, i: 9699] avg mini-batch loss: 0.984
[epoch: 0, i: 9799] avg mini-batch loss: 0.948
[epoch: 0, i: 9899] avg mini-batch loss: 0.923
[epoch: 0, i: 9999] avg mini-batch loss: 1.014
[epoch: 0, i: 10099] avg mini-batch loss: 1.025
[epoch: 0, i: 10199] avg mini-batch loss: 0.893
[epoch: 0, i: 10299] avg mini-batch loss: 1.015
[epoch: 0, i: 10399] avg mini-batch loss: 1.039
[epoch: 0, i: 10499] avg mini-batch loss: 0.905
[epoch: 0, i: 10599] avg mini-batch loss: 1.042
[epoch: 0, i: 10699] avg mini-batch loss: 0.913
[epoch: 0, i: 10799] avg mini-batch loss: 1.058
[epoch: 0, i: 10899] avg mini-batch loss: 1.064

[epoch: 0, i: 10999] avg mini-batch loss: 0.971
[epoch: 0, i: 11099] avg mini-batch loss: 1.026
[epoch: 0, i: 11199] avg mini-batch loss: 0.916
[epoch: 0, i: 11299] avg mini-batch loss: 1.069
[epoch: 0, i: 11399] avg mini-batch loss: 0.976
[epoch: 0, i: 11499] avg mini-batch loss: 0.936
[epoch: 0, i: 11599] avg mini-batch loss: 0.983
[epoch: 0, i: 11699] avg mini-batch loss: 0.873
[epoch: 0, i: 11799] avg mini-batch loss: 0.919
[epoch: 0, i: 11899] avg mini-batch loss: 0.964
[epoch: 0, i: 11999] avg mini-batch loss: 0.939
[epoch: 0, i: 12099] avg mini-batch loss: 0.971
[epoch: 0, i: 12199] avg mini-batch loss: 1.021
[epoch: 0, i: 12299] avg mini-batch loss: 0.893
[epoch: 0, i: 12399] avg mini-batch loss: 1.017
[epoch: 0, i: 12499] avg mini-batch loss: 0.876
[epoch: 1, i: 99] avg mini-batch loss: 0.951
[epoch: 1, i: 199] avg mini-batch loss: 0.841
[epoch: 1, i: 299] avg mini-batch loss: 0.783
[epoch: 1, i: 399] avg mini-batch loss: 0.955
[epoch: 1, i: 499] avg mini-batch loss: 0.835
[epoch: 1, i: 599] avg mini-batch loss: 0.856
[epoch: 1, i: 699] avg mini-batch loss: 0.914
[epoch: 1, i: 799] avg mini-batch loss: 0.787
[epoch: 1, i: 899] avg mini-batch loss: 0.918
[epoch: 1, i: 999] avg mini-batch loss: 0.883
[epoch: 1, i: 1099] avg mini-batch loss: 0.865
[epoch: 1, i: 1199] avg mini-batch loss: 0.788
[epoch: 1, i: 1299] avg mini-batch loss: 0.869
[epoch: 1, i: 1399] avg mini-batch loss: 0.879
[epoch: 1, i: 1499] avg mini-batch loss: 0.922
[epoch: 1, i: 1599] avg mini-batch loss: 0.798
[epoch: 1, i: 1699] avg mini-batch loss: 0.880
[epoch: 1, i: 1799] avg mini-batch loss: 0.773
[epoch: 1, i: 1899] avg mini-batch loss: 0.827
[epoch: 1, i: 1999] avg mini-batch loss: 0.817
[epoch: 1, i: 2099] avg mini-batch loss: 0.831
[epoch: 1, i: 2199] avg mini-batch loss: 0.827
[epoch: 1, i: 2299] avg mini-batch loss: 0.836
[epoch: 1, i: 2399] avg mini-batch loss: 0.810
[epoch: 1, i: 2499] avg mini-batch loss: 0.822
[epoch: 1, i: 2599] avg mini-batch loss: 0.955
[epoch: 1, i: 2699] avg mini-batch loss: 0.774
[epoch: 1, i: 2799] avg mini-batch loss: 0.899
[epoch: 1, i: 2899] avg mini-batch loss: 0.792
[epoch: 1, i: 2999] avg mini-batch loss: 0.839
[epoch: 1, i: 3099] avg mini-batch loss: 0.849
[epoch: 1, i: 3199] avg mini-batch loss: 0.779

[epoch: 1, i: 3299] avg mini-batch loss: 0.884
[epoch: 1, i: 3399] avg mini-batch loss: 0.906
[epoch: 1, i: 3499] avg mini-batch loss: 0.861
[epoch: 1, i: 3599] avg mini-batch loss: 0.922
[epoch: 1, i: 3699] avg mini-batch loss: 0.867
[epoch: 1, i: 3799] avg mini-batch loss: 0.807
[epoch: 1, i: 3899] avg mini-batch loss: 0.889
[epoch: 1, i: 3999] avg mini-batch loss: 0.743
[epoch: 1, i: 4099] avg mini-batch loss: 0.766
[epoch: 1, i: 4199] avg mini-batch loss: 0.795
[epoch: 1, i: 4299] avg mini-batch loss: 0.756
[epoch: 1, i: 4399] avg mini-batch loss: 0.817
[epoch: 1, i: 4499] avg mini-batch loss: 0.803
[epoch: 1, i: 4599] avg mini-batch loss: 0.846
[epoch: 1, i: 4699] avg mini-batch loss: 0.676
[epoch: 1, i: 4799] avg mini-batch loss: 0.791
[epoch: 1, i: 4899] avg mini-batch loss: 0.834
[epoch: 1, i: 4999] avg mini-batch loss: 0.882
[epoch: 1, i: 5099] avg mini-batch loss: 0.846
[epoch: 1, i: 5199] avg mini-batch loss: 0.908
[epoch: 1, i: 5299] avg mini-batch loss: 0.815
[epoch: 1, i: 5399] avg mini-batch loss: 0.779
[epoch: 1, i: 5499] avg mini-batch loss: 0.776
[epoch: 1, i: 5599] avg mini-batch loss: 0.824
[epoch: 1, i: 5699] avg mini-batch loss: 0.802
[epoch: 1, i: 5799] avg mini-batch loss: 0.799
[epoch: 1, i: 5899] avg mini-batch loss: 0.793
[epoch: 1, i: 5999] avg mini-batch loss: 0.789
[epoch: 1, i: 6099] avg mini-batch loss: 0.831
[epoch: 1, i: 6199] avg mini-batch loss: 0.800
[epoch: 1, i: 6299] avg mini-batch loss: 0.809
[epoch: 1, i: 6399] avg mini-batch loss: 0.783
[epoch: 1, i: 6499] avg mini-batch loss: 0.816
[epoch: 1, i: 6599] avg mini-batch loss: 0.832
[epoch: 1, i: 6699] avg mini-batch loss: 0.830
[epoch: 1, i: 6799] avg mini-batch loss: 0.782
[epoch: 1, i: 6899] avg mini-batch loss: 0.835
[epoch: 1, i: 6999] avg mini-batch loss: 0.770
[epoch: 1, i: 7099] avg mini-batch loss: 0.744
[epoch: 1, i: 7199] avg mini-batch loss: 0.831
[epoch: 1, i: 7299] avg mini-batch loss: 0.831
[epoch: 1, i: 7399] avg mini-batch loss: 0.822
[epoch: 1, i: 7499] avg mini-batch loss: 0.838
[epoch: 1, i: 7599] avg mini-batch loss: 0.827
[epoch: 1, i: 7699] avg mini-batch loss: 0.756
[epoch: 1, i: 7799] avg mini-batch loss: 0.815
[epoch: 1, i: 7899] avg mini-batch loss: 0.783
[epoch: 1, i: 7999] avg mini-batch loss: 0.811

[epoch: 1, i: 8099] avg mini-batch loss: 0.860
[epoch: 1, i: 8199] avg mini-batch loss: 0.703
[epoch: 1, i: 8299] avg mini-batch loss: 0.754
[epoch: 1, i: 8399] avg mini-batch loss: 0.688
[epoch: 1, i: 8499] avg mini-batch loss: 0.809
[epoch: 1, i: 8599] avg mini-batch loss: 0.840
[epoch: 1, i: 8699] avg mini-batch loss: 0.786
[epoch: 1, i: 8799] avg mini-batch loss: 0.915
[epoch: 1, i: 8899] avg mini-batch loss: 0.702
[epoch: 1, i: 8999] avg mini-batch loss: 0.781
[epoch: 1, i: 9099] avg mini-batch loss: 0.731
[epoch: 1, i: 9199] avg mini-batch loss: 0.794
[epoch: 1, i: 9299] avg mini-batch loss: 0.825
[epoch: 1, i: 9399] avg mini-batch loss: 0.729
[epoch: 1, i: 9499] avg mini-batch loss: 0.832
[epoch: 1, i: 9599] avg mini-batch loss: 0.746
[epoch: 1, i: 9699] avg mini-batch loss: 0.780
[epoch: 1, i: 9799] avg mini-batch loss: 0.696
[epoch: 1, i: 9899] avg mini-batch loss: 0.794
[epoch: 1, i: 9999] avg mini-batch loss: 0.700
[epoch: 1, i: 10099] avg mini-batch loss: 0.787
[epoch: 1, i: 10199] avg mini-batch loss: 0.756
[epoch: 1, i: 10299] avg mini-batch loss: 0.700
[epoch: 1, i: 10399] avg mini-batch loss: 0.849
[epoch: 1, i: 10499] avg mini-batch loss: 0.783
[epoch: 1, i: 10599] avg mini-batch loss: 0.739
[epoch: 1, i: 10699] avg mini-batch loss: 0.719
[epoch: 1, i: 10799] avg mini-batch loss: 0.697
[epoch: 1, i: 10899] avg mini-batch loss: 0.847
[epoch: 1, i: 10999] avg mini-batch loss: 0.849
[epoch: 1, i: 11099] avg mini-batch loss: 0.740
[epoch: 1, i: 11199] avg mini-batch loss: 0.751
[epoch: 1, i: 11299] avg mini-batch loss: 0.732
[epoch: 1, i: 11399] avg mini-batch loss: 0.772
[epoch: 1, i: 11499] avg mini-batch loss: 0.850
[epoch: 1, i: 11599] avg mini-batch loss: 0.675
[epoch: 1, i: 11699] avg mini-batch loss: 0.770
[epoch: 1, i: 11799] avg mini-batch loss: 0.761
[epoch: 1, i: 11899] avg mini-batch loss: 0.754
[epoch: 1, i: 11999] avg mini-batch loss: 0.666
[epoch: 1, i: 12099] avg mini-batch loss: 0.787
[epoch: 1, i: 12199] avg mini-batch loss: 0.711
[epoch: 1, i: 12299] avg mini-batch loss: 0.822
[epoch: 1, i: 12399] avg mini-batch loss: 0.853
[epoch: 1, i: 12499] avg mini-batch loss: 0.699
[epoch: 2, i: 99] avg mini-batch loss: 0.636
[epoch: 2, i: 199] avg mini-batch loss: 0.697
[epoch: 2, i: 299] avg mini-batch loss: 0.633

[epoch: 2, i: 399] avg mini-batch loss: 0.642
[epoch: 2, i: 499] avg mini-batch loss: 0.608
[epoch: 2, i: 599] avg mini-batch loss: 0.677
[epoch: 2, i: 699] avg mini-batch loss: 0.571
[epoch: 2, i: 799] avg mini-batch loss: 0.584
[epoch: 2, i: 899] avg mini-batch loss: 0.613
[epoch: 2, i: 999] avg mini-batch loss: 0.674
[epoch: 2, i: 1099] avg mini-batch loss: 0.624
[epoch: 2, i: 1199] avg mini-batch loss: 0.500
[epoch: 2, i: 1299] avg mini-batch loss: 0.585
[epoch: 2, i: 1399] avg mini-batch loss: 0.692
[epoch: 2, i: 1499] avg mini-batch loss: 0.643
[epoch: 2, i: 1599] avg mini-batch loss: 0.557
[epoch: 2, i: 1699] avg mini-batch loss: 0.672
[epoch: 2, i: 1799] avg mini-batch loss: 0.732
[epoch: 2, i: 1899] avg mini-batch loss: 0.646
[epoch: 2, i: 1999] avg mini-batch loss: 0.655
[epoch: 2, i: 2099] avg mini-batch loss: 0.601
[epoch: 2, i: 2199] avg mini-batch loss: 0.543
[epoch: 2, i: 2299] avg mini-batch loss: 0.550
[epoch: 2, i: 2399] avg mini-batch loss: 0.644
[epoch: 2, i: 2499] avg mini-batch loss: 0.609
[epoch: 2, i: 2599] avg mini-batch loss: 0.653
[epoch: 2, i: 2699] avg mini-batch loss: 0.584
[epoch: 2, i: 2799] avg mini-batch loss: 0.617
[epoch: 2, i: 2899] avg mini-batch loss: 0.602
[epoch: 2, i: 2999] avg mini-batch loss: 0.626
[epoch: 2, i: 3099] avg mini-batch loss: 0.623
[epoch: 2, i: 3199] avg mini-batch loss: 0.564
[epoch: 2, i: 3299] avg mini-batch loss: 0.603
[epoch: 2, i: 3399] avg mini-batch loss: 0.575
[epoch: 2, i: 3499] avg mini-batch loss: 0.471
[epoch: 2, i: 3599] avg mini-batch loss: 0.610
[epoch: 2, i: 3699] avg mini-batch loss: 0.681
[epoch: 2, i: 3799] avg mini-batch loss: 0.668
[epoch: 2, i: 3899] avg mini-batch loss: 0.571
[epoch: 2, i: 3999] avg mini-batch loss: 0.583
[epoch: 2, i: 4099] avg mini-batch loss: 0.576
[epoch: 2, i: 4199] avg mini-batch loss: 0.753
[epoch: 2, i: 4299] avg mini-batch loss: 0.663
[epoch: 2, i: 4399] avg mini-batch loss: 0.597
[epoch: 2, i: 4499] avg mini-batch loss: 0.633
[epoch: 2, i: 4599] avg mini-batch loss: 0.565
[epoch: 2, i: 4699] avg mini-batch loss: 0.636
[epoch: 2, i: 4799] avg mini-batch loss: 0.573
[epoch: 2, i: 4899] avg mini-batch loss: 0.625
[epoch: 2, i: 4999] avg mini-batch loss: 0.760
[epoch: 2, i: 5099] avg mini-batch loss: 0.604

[epoch: 2, i: 5199] avg mini-batch loss: 0.600
[epoch: 2, i: 5299] avg mini-batch loss: 0.652
[epoch: 2, i: 5399] avg mini-batch loss: 0.624
[epoch: 2, i: 5499] avg mini-batch loss: 0.671
[epoch: 2, i: 5599] avg mini-batch loss: 0.653
[epoch: 2, i: 5699] avg mini-batch loss: 0.676
[epoch: 2, i: 5799] avg mini-batch loss: 0.636
[epoch: 2, i: 5899] avg mini-batch loss: 0.632
[epoch: 2, i: 5999] avg mini-batch loss: 0.635
[epoch: 2, i: 6099] avg mini-batch loss: 0.644
[epoch: 2, i: 6199] avg mini-batch loss: 0.613
[epoch: 2, i: 6299] avg mini-batch loss: 0.680
[epoch: 2, i: 6399] avg mini-batch loss: 0.661
[epoch: 2, i: 6499] avg mini-batch loss: 0.622
[epoch: 2, i: 6599] avg mini-batch loss: 0.575
[epoch: 2, i: 6699] avg mini-batch loss: 0.686
[epoch: 2, i: 6799] avg mini-batch loss: 0.669
[epoch: 2, i: 6899] avg mini-batch loss: 0.648
[epoch: 2, i: 6999] avg mini-batch loss: 0.675
[epoch: 2, i: 7099] avg mini-batch loss: 0.673
[epoch: 2, i: 7199] avg mini-batch loss: 0.602
[epoch: 2, i: 7299] avg mini-batch loss: 0.568
[epoch: 2, i: 7399] avg mini-batch loss: 0.550
[epoch: 2, i: 7499] avg mini-batch loss: 0.611
[epoch: 2, i: 7599] avg mini-batch loss: 0.613
[epoch: 2, i: 7699] avg mini-batch loss: 0.699
[epoch: 2, i: 7799] avg mini-batch loss: 0.596
[epoch: 2, i: 7899] avg mini-batch loss: 0.622
[epoch: 2, i: 7999] avg mini-batch loss: 0.635
[epoch: 2, i: 8099] avg mini-batch loss: 0.593
[epoch: 2, i: 8199] avg mini-batch loss: 0.588
[epoch: 2, i: 8299] avg mini-batch loss: 0.649
[epoch: 2, i: 8399] avg mini-batch loss: 0.674
[epoch: 2, i: 8499] avg mini-batch loss: 0.556
[epoch: 2, i: 8599] avg mini-batch loss: 0.618
[epoch: 2, i: 8699] avg mini-batch loss: 0.667
[epoch: 2, i: 8799] avg mini-batch loss: 0.508
[epoch: 2, i: 8899] avg mini-batch loss: 0.667
[epoch: 2, i: 8999] avg mini-batch loss: 0.615
[epoch: 2, i: 9099] avg mini-batch loss: 0.616
[epoch: 2, i: 9199] avg mini-batch loss: 0.616
[epoch: 2, i: 9299] avg mini-batch loss: 0.663
[epoch: 2, i: 9399] avg mini-batch loss: 0.659
[epoch: 2, i: 9499] avg mini-batch loss: 0.535
[epoch: 2, i: 9599] avg mini-batch loss: 0.657
[epoch: 2, i: 9699] avg mini-batch loss: 0.615
[epoch: 2, i: 9799] avg mini-batch loss: 0.583
[epoch: 2, i: 9899] avg mini-batch loss: 0.662

[epoch: 2, i: 9999] avg mini-batch loss: 0.657
[epoch: 2, i: 10099] avg mini-batch loss: 0.606
[epoch: 2, i: 10199] avg mini-batch loss: 0.606
[epoch: 2, i: 10299] avg mini-batch loss: 0.602
[epoch: 2, i: 10399] avg mini-batch loss: 0.625
[epoch: 2, i: 10499] avg mini-batch loss: 0.650
[epoch: 2, i: 10599] avg mini-batch loss: 0.646
[epoch: 2, i: 10699] avg mini-batch loss: 0.602
[epoch: 2, i: 10799] avg mini-batch loss: 0.639
[epoch: 2, i: 10899] avg mini-batch loss: 0.652
[epoch: 2, i: 10999] avg mini-batch loss: 0.589
[epoch: 2, i: 11099] avg mini-batch loss: 0.609
[epoch: 2, i: 11199] avg mini-batch loss: 0.668
[epoch: 2, i: 11299] avg mini-batch loss: 0.638
[epoch: 2, i: 11399] avg mini-batch loss: 0.589
[epoch: 2, i: 11499] avg mini-batch loss: 0.671
[epoch: 2, i: 11599] avg mini-batch loss: 0.625
[epoch: 2, i: 11699] avg mini-batch loss: 0.628
[epoch: 2, i: 11799] avg mini-batch loss: 0.639
[epoch: 2, i: 11899] avg mini-batch loss: 0.578
[epoch: 2, i: 11999] avg mini-batch loss: 0.640
[epoch: 2, i: 12099] avg mini-batch loss: 0.576
[epoch: 2, i: 12199] avg mini-batch loss: 0.558
[epoch: 2, i: 12299] avg mini-batch loss: 0.595
[epoch: 2, i: 12399] avg mini-batch loss: 0.613
[epoch: 2, i: 12499] avg mini-batch loss: 0.622
[epoch: 3, i: 99] avg mini-batch loss: 0.431
[epoch: 3, i: 199] avg mini-batch loss: 0.498
[epoch: 3, i: 299] avg mini-batch loss: 0.437
[epoch: 3, i: 399] avg mini-batch loss: 0.411
[epoch: 3, i: 499] avg mini-batch loss: 0.420
[epoch: 3, i: 599] avg mini-batch loss: 0.463
[epoch: 3, i: 699] avg mini-batch loss: 0.448
[epoch: 3, i: 799] avg mini-batch loss: 0.489
[epoch: 3, i: 899] avg mini-batch loss: 0.475
[epoch: 3, i: 999] avg mini-batch loss: 0.411
[epoch: 3, i: 1099] avg mini-batch loss: 0.518
[epoch: 3, i: 1199] avg mini-batch loss: 0.385
[epoch: 3, i: 1299] avg mini-batch loss: 0.506
[epoch: 3, i: 1399] avg mini-batch loss: 0.454
[epoch: 3, i: 1499] avg mini-batch loss: 0.468
[epoch: 3, i: 1599] avg mini-batch loss: 0.504
[epoch: 3, i: 1699] avg mini-batch loss: 0.408
[epoch: 3, i: 1799] avg mini-batch loss: 0.487
[epoch: 3, i: 1899] avg mini-batch loss: 0.393
[epoch: 3, i: 1999] avg mini-batch loss: 0.492
[epoch: 3, i: 2099] avg mini-batch loss: 0.442
[epoch: 3, i: 2199] avg mini-batch loss: 0.377

[epoch: 3, i: 2299] avg mini-batch loss: 0.483
[epoch: 3, i: 2399] avg mini-batch loss: 0.430
[epoch: 3, i: 2499] avg mini-batch loss: 0.429
[epoch: 3, i: 2599] avg mini-batch loss: 0.426
[epoch: 3, i: 2699] avg mini-batch loss: 0.460
[epoch: 3, i: 2799] avg mini-batch loss: 0.444
[epoch: 3, i: 2899] avg mini-batch loss: 0.501
[epoch: 3, i: 2999] avg mini-batch loss: 0.426
[epoch: 3, i: 3099] avg mini-batch loss: 0.514
[epoch: 3, i: 3199] avg mini-batch loss: 0.385
[epoch: 3, i: 3299] avg mini-batch loss: 0.455
[epoch: 3, i: 3399] avg mini-batch loss: 0.590
[epoch: 3, i: 3499] avg mini-batch loss: 0.443
[epoch: 3, i: 3599] avg mini-batch loss: 0.397
[epoch: 3, i: 3699] avg mini-batch loss: 0.461
[epoch: 3, i: 3799] avg mini-batch loss: 0.519
[epoch: 3, i: 3899] avg mini-batch loss: 0.431
[epoch: 3, i: 3999] avg mini-batch loss: 0.554
[epoch: 3, i: 4099] avg mini-batch loss: 0.451
[epoch: 3, i: 4199] avg mini-batch loss: 0.503
[epoch: 3, i: 4299] avg mini-batch loss: 0.473
[epoch: 3, i: 4399] avg mini-batch loss: 0.569
[epoch: 3, i: 4499] avg mini-batch loss: 0.488
[epoch: 3, i: 4599] avg mini-batch loss: 0.436
[epoch: 3, i: 4699] avg mini-batch loss: 0.498
[epoch: 3, i: 4799] avg mini-batch loss: 0.410
[epoch: 3, i: 4899] avg mini-batch loss: 0.420
[epoch: 3, i: 4999] avg mini-batch loss: 0.468
[epoch: 3, i: 5099] avg mini-batch loss: 0.489
[epoch: 3, i: 5199] avg mini-batch loss: 0.518
[epoch: 3, i: 5299] avg mini-batch loss: 0.486
[epoch: 3, i: 5399] avg mini-batch loss: 0.558
[epoch: 3, i: 5499] avg mini-batch loss: 0.479
[epoch: 3, i: 5599] avg mini-batch loss: 0.452
[epoch: 3, i: 5699] avg mini-batch loss: 0.442
[epoch: 3, i: 5799] avg mini-batch loss: 0.497
[epoch: 3, i: 5899] avg mini-batch loss: 0.472
[epoch: 3, i: 5999] avg mini-batch loss: 0.463
[epoch: 3, i: 6099] avg mini-batch loss: 0.552
[epoch: 3, i: 6199] avg mini-batch loss: 0.498
[epoch: 3, i: 6299] avg mini-batch loss: 0.438
[epoch: 3, i: 6399] avg mini-batch loss: 0.496
[epoch: 3, i: 6499] avg mini-batch loss: 0.507
[epoch: 3, i: 6599] avg mini-batch loss: 0.510
[epoch: 3, i: 6699] avg mini-batch loss: 0.510
[epoch: 3, i: 6799] avg mini-batch loss: 0.506
[epoch: 3, i: 6899] avg mini-batch loss: 0.469
[epoch: 3, i: 6999] avg mini-batch loss: 0.539

[epoch: 3, i: 7099] avg mini-batch loss: 0.501
[epoch: 3, i: 7199] avg mini-batch loss: 0.450
[epoch: 3, i: 7299] avg mini-batch loss: 0.466
[epoch: 3, i: 7399] avg mini-batch loss: 0.396
[epoch: 3, i: 7499] avg mini-batch loss: 0.471
[epoch: 3, i: 7599] avg mini-batch loss: 0.390
[epoch: 3, i: 7699] avg mini-batch loss: 0.470
[epoch: 3, i: 7799] avg mini-batch loss: 0.490
[epoch: 3, i: 7899] avg mini-batch loss: 0.407
[epoch: 3, i: 7999] avg mini-batch loss: 0.483
[epoch: 3, i: 8099] avg mini-batch loss: 0.507
[epoch: 3, i: 8199] avg mini-batch loss: 0.444
[epoch: 3, i: 8299] avg mini-batch loss: 0.507
[epoch: 3, i: 8399] avg mini-batch loss: 0.481
[epoch: 3, i: 8499] avg mini-batch loss: 0.486
[epoch: 3, i: 8599] avg mini-batch loss: 0.498
[epoch: 3, i: 8699] avg mini-batch loss: 0.476
[epoch: 3, i: 8799] avg mini-batch loss: 0.525
[epoch: 3, i: 8899] avg mini-batch loss: 0.391
[epoch: 3, i: 8999] avg mini-batch loss: 0.380
[epoch: 3, i: 9099] avg mini-batch loss: 0.463
[epoch: 3, i: 9199] avg mini-batch loss: 0.533
[epoch: 3, i: 9299] avg mini-batch loss: 0.458
[epoch: 3, i: 9399] avg mini-batch loss: 0.421
[epoch: 3, i: 9499] avg mini-batch loss: 0.419
[epoch: 3, i: 9599] avg mini-batch loss: 0.496
[epoch: 3, i: 9699] avg mini-batch loss: 0.455
[epoch: 3, i: 9799] avg mini-batch loss: 0.487
[epoch: 3, i: 9899] avg mini-batch loss: 0.475
[epoch: 3, i: 9999] avg mini-batch loss: 0.431
[epoch: 3, i: 10099] avg mini-batch loss: 0.450
[epoch: 3, i: 10199] avg mini-batch loss: 0.515
[epoch: 3, i: 10299] avg mini-batch loss: 0.421
[epoch: 3, i: 10399] avg mini-batch loss: 0.565
[epoch: 3, i: 10499] avg mini-batch loss: 0.510
[epoch: 3, i: 10599] avg mini-batch loss: 0.446
[epoch: 3, i: 10699] avg mini-batch loss: 0.408
[epoch: 3, i: 10799] avg mini-batch loss: 0.555
[epoch: 3, i: 10899] avg mini-batch loss: 0.456
[epoch: 3, i: 10999] avg mini-batch loss: 0.505
[epoch: 3, i: 11099] avg mini-batch loss: 0.456
[epoch: 3, i: 11199] avg mini-batch loss: 0.529
[epoch: 3, i: 11299] avg mini-batch loss: 0.464
[epoch: 3, i: 11399] avg mini-batch loss: 0.538
[epoch: 3, i: 11499] avg mini-batch loss: 0.482
[epoch: 3, i: 11599] avg mini-batch loss: 0.482
[epoch: 3, i: 11699] avg mini-batch loss: 0.466
[epoch: 3, i: 11799] avg mini-batch loss: 0.468

[epoch: 3, i: 11899] avg mini-batch loss: 0.568
[epoch: 3, i: 11999] avg mini-batch loss: 0.496
[epoch: 3, i: 12099] avg mini-batch loss: 0.440
[epoch: 3, i: 12199] avg mini-batch loss: 0.479
[epoch: 3, i: 12299] avg mini-batch loss: 0.505
[epoch: 3, i: 12399] avg mini-batch loss: 0.484
[epoch: 3, i: 12499] avg mini-batch loss: 0.462
[epoch: 4, i: 99] avg mini-batch loss: 0.246
[epoch: 4, i: 199] avg mini-batch loss: 0.322
[epoch: 4, i: 299] avg mini-batch loss: 0.259
[epoch: 4, i: 399] avg mini-batch loss: 0.300
[epoch: 4, i: 499] avg mini-batch loss: 0.264
[epoch: 4, i: 599] avg mini-batch loss: 0.374
[epoch: 4, i: 699] avg mini-batch loss: 0.259
[epoch: 4, i: 799] avg mini-batch loss: 0.321
[epoch: 4, i: 899] avg mini-batch loss: 0.213
[epoch: 4, i: 999] avg mini-batch loss: 0.241
[epoch: 4, i: 1099] avg mini-batch loss: 0.297
[epoch: 4, i: 1199] avg mini-batch loss: 0.310
[epoch: 4, i: 1299] avg mini-batch loss: 0.332
[epoch: 4, i: 1399] avg mini-batch loss: 0.347
[epoch: 4, i: 1499] avg mini-batch loss: 0.331
[epoch: 4, i: 1599] avg mini-batch loss: 0.334
[epoch: 4, i: 1699] avg mini-batch loss: 0.311
[epoch: 4, i: 1799] avg mini-batch loss: 0.273
[epoch: 4, i: 1899] avg mini-batch loss: 0.225
[epoch: 4, i: 1999] avg mini-batch loss: 0.264
[epoch: 4, i: 2099] avg mini-batch loss: 0.262
[epoch: 4, i: 2199] avg mini-batch loss: 0.279
[epoch: 4, i: 2299] avg mini-batch loss: 0.302
[epoch: 4, i: 2399] avg mini-batch loss: 0.364
[epoch: 4, i: 2499] avg mini-batch loss: 0.353
[epoch: 4, i: 2599] avg mini-batch loss: 0.321
[epoch: 4, i: 2699] avg mini-batch loss: 0.312
[epoch: 4, i: 2799] avg mini-batch loss: 0.319
[epoch: 4, i: 2899] avg mini-batch loss: 0.302
[epoch: 4, i: 2999] avg mini-batch loss: 0.266
[epoch: 4, i: 3099] avg mini-batch loss: 0.311
[epoch: 4, i: 3199] avg mini-batch loss: 0.337
[epoch: 4, i: 3299] avg mini-batch loss: 0.315
[epoch: 4, i: 3399] avg mini-batch loss: 0.359
[epoch: 4, i: 3499] avg mini-batch loss: 0.280
[epoch: 4, i: 3599] avg mini-batch loss: 0.267
[epoch: 4, i: 3699] avg mini-batch loss: 0.325
[epoch: 4, i: 3799] avg mini-batch loss: 0.345
[epoch: 4, i: 3899] avg mini-batch loss: 0.279
[epoch: 4, i: 3999] avg mini-batch loss: 0.381
[epoch: 4, i: 4099] avg mini-batch loss: 0.342

[epoch: 4, i: 4199] avg mini-batch loss: 0.361
[epoch: 4, i: 4299] avg mini-batch loss: 0.361
[epoch: 4, i: 4399] avg mini-batch loss: 0.404
[epoch: 4, i: 4499] avg mini-batch loss: 0.342
[epoch: 4, i: 4599] avg mini-batch loss: 0.339
[epoch: 4, i: 4699] avg mini-batch loss: 0.254
[epoch: 4, i: 4799] avg mini-batch loss: 0.320
[epoch: 4, i: 4899] avg mini-batch loss: 0.368
[epoch: 4, i: 4999] avg mini-batch loss: 0.312
[epoch: 4, i: 5099] avg mini-batch loss: 0.299
[epoch: 4, i: 5199] avg mini-batch loss: 0.337
[epoch: 4, i: 5299] avg mini-batch loss: 0.313
[epoch: 4, i: 5399] avg mini-batch loss: 0.345
[epoch: 4, i: 5499] avg mini-batch loss: 0.371
[epoch: 4, i: 5599] avg mini-batch loss: 0.280
[epoch: 4, i: 5699] avg mini-batch loss: 0.263
[epoch: 4, i: 5799] avg mini-batch loss: 0.320
[epoch: 4, i: 5899] avg mini-batch loss: 0.375
[epoch: 4, i: 5999] avg mini-batch loss: 0.309
[epoch: 4, i: 6099] avg mini-batch loss: 0.362
[epoch: 4, i: 6199] avg mini-batch loss: 0.360
[epoch: 4, i: 6299] avg mini-batch loss: 0.325
[epoch: 4, i: 6399] avg mini-batch loss: 0.326
[epoch: 4, i: 6499] avg mini-batch loss: 0.322
[epoch: 4, i: 6599] avg mini-batch loss: 0.352
[epoch: 4, i: 6699] avg mini-batch loss: 0.420
[epoch: 4, i: 6799] avg mini-batch loss: 0.336
[epoch: 4, i: 6899] avg mini-batch loss: 0.346
[epoch: 4, i: 6999] avg mini-batch loss: 0.359
[epoch: 4, i: 7099] avg mini-batch loss: 0.310
[epoch: 4, i: 7199] avg mini-batch loss: 0.358
[epoch: 4, i: 7299] avg mini-batch loss: 0.405
[epoch: 4, i: 7399] avg mini-batch loss: 0.304
[epoch: 4, i: 7499] avg mini-batch loss: 0.285
[epoch: 4, i: 7599] avg mini-batch loss: 0.337
[epoch: 4, i: 7699] avg mini-batch loss: 0.343
[epoch: 4, i: 7799] avg mini-batch loss: 0.238
[epoch: 4, i: 7899] avg mini-batch loss: 0.306
[epoch: 4, i: 7999] avg mini-batch loss: 0.398
[epoch: 4, i: 8099] avg mini-batch loss: 0.308
[epoch: 4, i: 8199] avg mini-batch loss: 0.303
[epoch: 4, i: 8299] avg mini-batch loss: 0.415
[epoch: 4, i: 8399] avg mini-batch loss: 0.375
[epoch: 4, i: 8499] avg mini-batch loss: 0.306
[epoch: 4, i: 8599] avg mini-batch loss: 0.339
[epoch: 4, i: 8699] avg mini-batch loss: 0.303
[epoch: 4, i: 8799] avg mini-batch loss: 0.338
[epoch: 4, i: 8899] avg mini-batch loss: 0.421

[epoch: 4, i: 8999] avg mini-batch loss: 0.346
[epoch: 4, i: 9099] avg mini-batch loss: 0.295
[epoch: 4, i: 9199] avg mini-batch loss: 0.360
[epoch: 4, i: 9299] avg mini-batch loss: 0.373
[epoch: 4, i: 9399] avg mini-batch loss: 0.392
[epoch: 4, i: 9499] avg mini-batch loss: 0.311
[epoch: 4, i: 9599] avg mini-batch loss: 0.321
[epoch: 4, i: 9699] avg mini-batch loss: 0.344
[epoch: 4, i: 9799] avg mini-batch loss: 0.382
[epoch: 4, i: 9899] avg mini-batch loss: 0.324
[epoch: 4, i: 9999] avg mini-batch loss: 0.344
[epoch: 4, i: 10099] avg mini-batch loss: 0.349
[epoch: 4, i: 10199] avg mini-batch loss: 0.332
[epoch: 4, i: 10299] avg mini-batch loss: 0.376
[epoch: 4, i: 10399] avg mini-batch loss: 0.388
[epoch: 4, i: 10499] avg mini-batch loss: 0.355
[epoch: 4, i: 10599] avg mini-batch loss: 0.376
[epoch: 4, i: 10699] avg mini-batch loss: 0.295
[epoch: 4, i: 10799] avg mini-batch loss: 0.377
[epoch: 4, i: 10899] avg mini-batch loss: 0.378
[epoch: 4, i: 10999] avg mini-batch loss: 0.403
[epoch: 4, i: 11099] avg mini-batch loss: 0.381
[epoch: 4, i: 11199] avg mini-batch loss: 0.352
[epoch: 4, i: 11299] avg mini-batch loss: 0.359
[epoch: 4, i: 11399] avg mini-batch loss: 0.380
[epoch: 4, i: 11499] avg mini-batch loss: 0.400
[epoch: 4, i: 11599] avg mini-batch loss: 0.350
[epoch: 4, i: 11699] avg mini-batch loss: 0.356
[epoch: 4, i: 11799] avg mini-batch loss: 0.293
[epoch: 4, i: 11899] avg mini-batch loss: 0.405
[epoch: 4, i: 11999] avg mini-batch loss: 0.339
[epoch: 4, i: 12099] avg mini-batch loss: 0.334
[epoch: 4, i: 12199] avg mini-batch loss: 0.328
[epoch: 4, i: 12299] avg mini-batch loss: 0.379
[epoch: 4, i: 12399] avg mini-batch loss: 0.394
[epoch: 4, i: 12499] avg mini-batch loss: 0.391
[epoch: 5, i: 99] avg mini-batch loss: 0.212
[epoch: 5, i: 199] avg mini-batch loss: 0.224
[epoch: 5, i: 299] avg mini-batch loss: 0.150
[epoch: 5, i: 399] avg mini-batch loss: 0.166
[epoch: 5, i: 499] avg mini-batch loss: 0.168
[epoch: 5, i: 599] avg mini-batch loss: 0.130
[epoch: 5, i: 699] avg mini-batch loss: 0.160
[epoch: 5, i: 799] avg mini-batch loss: 0.135
[epoch: 5, i: 899] avg mini-batch loss: 0.233
[epoch: 5, i: 999] avg mini-batch loss: 0.156
[epoch: 5, i: 1099] avg mini-batch loss: 0.162
[epoch: 5, i: 1199] avg mini-batch loss: 0.195

[epoch: 5, i: 1299] avg mini-batch loss: 0.153
[epoch: 5, i: 1399] avg mini-batch loss: 0.171
[epoch: 5, i: 1499] avg mini-batch loss: 0.196
[epoch: 5, i: 1599] avg mini-batch loss: 0.171
[epoch: 5, i: 1699] avg mini-batch loss: 0.169
[epoch: 5, i: 1799] avg mini-batch loss: 0.192
[epoch: 5, i: 1899] avg mini-batch loss: 0.186
[epoch: 5, i: 1999] avg mini-batch loss: 0.179
[epoch: 5, i: 2099] avg mini-batch loss: 0.186
[epoch: 5, i: 2199] avg mini-batch loss: 0.173
[epoch: 5, i: 2299] avg mini-batch loss: 0.187
[epoch: 5, i: 2399] avg mini-batch loss: 0.150
[epoch: 5, i: 2499] avg mini-batch loss: 0.134
[epoch: 5, i: 2599] avg mini-batch loss: 0.206
[epoch: 5, i: 2699] avg mini-batch loss: 0.170
[epoch: 5, i: 2799] avg mini-batch loss: 0.174
[epoch: 5, i: 2899] avg mini-batch loss: 0.254
[epoch: 5, i: 2999] avg mini-batch loss: 0.194
[epoch: 5, i: 3099] avg mini-batch loss: 0.223
[epoch: 5, i: 3199] avg mini-batch loss: 0.195
[epoch: 5, i: 3299] avg mini-batch loss: 0.217
[epoch: 5, i: 3399] avg mini-batch loss: 0.226
[epoch: 5, i: 3499] avg mini-batch loss: 0.179
[epoch: 5, i: 3599] avg mini-batch loss: 0.255
[epoch: 5, i: 3699] avg mini-batch loss: 0.202
[epoch: 5, i: 3799] avg mini-batch loss: 0.198
[epoch: 5, i: 3899] avg mini-batch loss: 0.234
[epoch: 5, i: 3999] avg mini-batch loss: 0.188
[epoch: 5, i: 4099] avg mini-batch loss: 0.208
[epoch: 5, i: 4199] avg mini-batch loss: 0.206
[epoch: 5, i: 4299] avg mini-batch loss: 0.178
[epoch: 5, i: 4399] avg mini-batch loss: 0.209
[epoch: 5, i: 4499] avg mini-batch loss: 0.219
[epoch: 5, i: 4599] avg mini-batch loss: 0.240
[epoch: 5, i: 4699] avg mini-batch loss: 0.181
[epoch: 5, i: 4799] avg mini-batch loss: 0.149
[epoch: 5, i: 4899] avg mini-batch loss: 0.157
[epoch: 5, i: 4999] avg mini-batch loss: 0.204
[epoch: 5, i: 5099] avg mini-batch loss: 0.214
[epoch: 5, i: 5199] avg mini-batch loss: 0.213
[epoch: 5, i: 5299] avg mini-batch loss: 0.237
[epoch: 5, i: 5399] avg mini-batch loss: 0.197
[epoch: 5, i: 5499] avg mini-batch loss: 0.195
[epoch: 5, i: 5599] avg mini-batch loss: 0.191
[epoch: 5, i: 5699] avg mini-batch loss: 0.221
[epoch: 5, i: 5799] avg mini-batch loss: 0.232
[epoch: 5, i: 5899] avg mini-batch loss: 0.209
[epoch: 5, i: 5999] avg mini-batch loss: 0.262

[epoch: 5, i: 6099] avg mini-batch loss: 0.300
[epoch: 5, i: 6199] avg mini-batch loss: 0.272
[epoch: 5, i: 6299] avg mini-batch loss: 0.151
[epoch: 5, i: 6399] avg mini-batch loss: 0.237
[epoch: 5, i: 6499] avg mini-batch loss: 0.213
[epoch: 5, i: 6599] avg mini-batch loss: 0.257
[epoch: 5, i: 6699] avg mini-batch loss: 0.269
[epoch: 5, i: 6799] avg mini-batch loss: 0.202
[epoch: 5, i: 6899] avg mini-batch loss: 0.187
[epoch: 5, i: 6999] avg mini-batch loss: 0.191
[epoch: 5, i: 7099] avg mini-batch loss: 0.237
[epoch: 5, i: 7199] avg mini-batch loss: 0.206
[epoch: 5, i: 7299] avg mini-batch loss: 0.222
[epoch: 5, i: 7399] avg mini-batch loss: 0.228
[epoch: 5, i: 7499] avg mini-batch loss: 0.256
[epoch: 5, i: 7599] avg mini-batch loss: 0.221
[epoch: 5, i: 7699] avg mini-batch loss: 0.198
[epoch: 5, i: 7799] avg mini-batch loss: 0.225
[epoch: 5, i: 7899] avg mini-batch loss: 0.254
[epoch: 5, i: 7999] avg mini-batch loss: 0.186
[epoch: 5, i: 8099] avg mini-batch loss: 0.227
[epoch: 5, i: 8199] avg mini-batch loss: 0.191
[epoch: 5, i: 8299] avg mini-batch loss: 0.217
[epoch: 5, i: 8399] avg mini-batch loss: 0.225
[epoch: 5, i: 8499] avg mini-batch loss: 0.235
[epoch: 5, i: 8599] avg mini-batch loss: 0.204
[epoch: 5, i: 8699] avg mini-batch loss: 0.314
[epoch: 5, i: 8799] avg mini-batch loss: 0.237
[epoch: 5, i: 8899] avg mini-batch loss: 0.250
[epoch: 5, i: 8999] avg mini-batch loss: 0.236
[epoch: 5, i: 9099] avg mini-batch loss: 0.228
[epoch: 5, i: 9199] avg mini-batch loss: 0.271
[epoch: 5, i: 9299] avg mini-batch loss: 0.218
[epoch: 5, i: 9399] avg mini-batch loss: 0.204
[epoch: 5, i: 9499] avg mini-batch loss: 0.230
[epoch: 5, i: 9599] avg mini-batch loss: 0.262
[epoch: 5, i: 9699] avg mini-batch loss: 0.228
[epoch: 5, i: 9799] avg mini-batch loss: 0.221
[epoch: 5, i: 9899] avg mini-batch loss: 0.246
[epoch: 5, i: 9999] avg mini-batch loss: 0.279
[epoch: 5, i: 10099] avg mini-batch loss: 0.367
[epoch: 5, i: 10199] avg mini-batch loss: 0.236
[epoch: 5, i: 10299] avg mini-batch loss: 0.281
[epoch: 5, i: 10399] avg mini-batch loss: 0.216
[epoch: 5, i: 10499] avg mini-batch loss: 0.215
[epoch: 5, i: 10599] avg mini-batch loss: 0.282
[epoch: 5, i: 10699] avg mini-batch loss: 0.348
[epoch: 5, i: 10799] avg mini-batch loss: 0.192

[epoch: 5, i: 10899] avg mini-batch loss: 0.212
[epoch: 5, i: 10999] avg mini-batch loss: 0.259
[epoch: 5, i: 11099] avg mini-batch loss: 0.214
[epoch: 5, i: 11199] avg mini-batch loss: 0.240
[epoch: 5, i: 11299] avg mini-batch loss: 0.327
[epoch: 5, i: 11399] avg mini-batch loss: 0.233
[epoch: 5, i: 11499] avg mini-batch loss: 0.247
[epoch: 5, i: 11599] avg mini-batch loss: 0.292
[epoch: 5, i: 11699] avg mini-batch loss: 0.217
[epoch: 5, i: 11799] avg mini-batch loss: 0.233
[epoch: 5, i: 11899] avg mini-batch loss: 0.253
[epoch: 5, i: 11999] avg mini-batch loss: 0.211
[epoch: 5, i: 12099] avg mini-batch loss: 0.243
[epoch: 5, i: 12199] avg mini-batch loss: 0.234
[epoch: 5, i: 12299] avg mini-batch loss: 0.235
[epoch: 5, i: 12399] avg mini-batch loss: 0.256
[epoch: 5, i: 12499] avg mini-batch loss: 0.192
[epoch: 6, i: 99] avg mini-batch loss: 0.130
[epoch: 6, i: 199] avg mini-batch loss: 0.111
[epoch: 6, i: 299] avg mini-batch loss: 0.093
[epoch: 6, i: 399] avg mini-batch loss: 0.131
[epoch: 6, i: 499] avg mini-batch loss: 0.067
[epoch: 6, i: 599] avg mini-batch loss: 0.055
[epoch: 6, i: 699] avg mini-batch loss: 0.093
[epoch: 6, i: 799] avg mini-batch loss: 0.118
[epoch: 6, i: 899] avg mini-batch loss: 0.133
[epoch: 6, i: 999] avg mini-batch loss: 0.098
[epoch: 6, i: 1099] avg mini-batch loss: 0.105
[epoch: 6, i: 1199] avg mini-batch loss: 0.106
[epoch: 6, i: 1299] avg mini-batch loss: 0.075
[epoch: 6, i: 1399] avg mini-batch loss: 0.090
[epoch: 6, i: 1499] avg mini-batch loss: 0.089
[epoch: 6, i: 1599] avg mini-batch loss: 0.100
[epoch: 6, i: 1699] avg mini-batch loss: 0.076
[epoch: 6, i: 1799] avg mini-batch loss: 0.067
[epoch: 6, i: 1899] avg mini-batch loss: 0.058
[epoch: 6, i: 1999] avg mini-batch loss: 0.123
[epoch: 6, i: 2099] avg mini-batch loss: 0.125
[epoch: 6, i: 2199] avg mini-batch loss: 0.105
[epoch: 6, i: 2299] avg mini-batch loss: 0.101
[epoch: 6, i: 2399] avg mini-batch loss: 0.128
[epoch: 6, i: 2499] avg mini-batch loss: 0.079
[epoch: 6, i: 2599] avg mini-batch loss: 0.086
[epoch: 6, i: 2699] avg mini-batch loss: 0.086
[epoch: 6, i: 2799] avg mini-batch loss: 0.091
[epoch: 6, i: 2899] avg mini-batch loss: 0.059
[epoch: 6, i: 2999] avg mini-batch loss: 0.152
[epoch: 6, i: 3099] avg mini-batch loss: 0.095

[epoch: 6, i: 3199] avg mini-batch loss: 0.129
[epoch: 6, i: 3299] avg mini-batch loss: 0.203
[epoch: 6, i: 3399] avg mini-batch loss: 0.114
[epoch: 6, i: 3499] avg mini-batch loss: 0.113
[epoch: 6, i: 3599] avg mini-batch loss: 0.121
[epoch: 6, i: 3699] avg mini-batch loss: 0.099
[epoch: 6, i: 3799] avg mini-batch loss: 0.145
[epoch: 6, i: 3899] avg mini-batch loss: 0.141
[epoch: 6, i: 3999] avg mini-batch loss: 0.093
[epoch: 6, i: 4099] avg mini-batch loss: 0.127
[epoch: 6, i: 4199] avg mini-batch loss: 0.159
[epoch: 6, i: 4299] avg mini-batch loss: 0.143
[epoch: 6, i: 4399] avg mini-batch loss: 0.112
[epoch: 6, i: 4499] avg mini-batch loss: 0.177
[epoch: 6, i: 4599] avg mini-batch loss: 0.115
[epoch: 6, i: 4699] avg mini-batch loss: 0.198
[epoch: 6, i: 4799] avg mini-batch loss: 0.120
[epoch: 6, i: 4899] avg mini-batch loss: 0.108
[epoch: 6, i: 4999] avg mini-batch loss: 0.109
[epoch: 6, i: 5099] avg mini-batch loss: 0.146
[epoch: 6, i: 5199] avg mini-batch loss: 0.140
[epoch: 6, i: 5299] avg mini-batch loss: 0.134
[epoch: 6, i: 5399] avg mini-batch loss: 0.121
[epoch: 6, i: 5499] avg mini-batch loss: 0.163
[epoch: 6, i: 5599] avg mini-batch loss: 0.148
[epoch: 6, i: 5699] avg mini-batch loss: 0.115
[epoch: 6, i: 5799] avg mini-batch loss: 0.141
[epoch: 6, i: 5899] avg mini-batch loss: 0.147
[epoch: 6, i: 5999] avg mini-batch loss: 0.146
[epoch: 6, i: 6099] avg mini-batch loss: 0.095
[epoch: 6, i: 6199] avg mini-batch loss: 0.081
[epoch: 6, i: 6299] avg mini-batch loss: 0.138
[epoch: 6, i: 6399] avg mini-batch loss: 0.128
[epoch: 6, i: 6499] avg mini-batch loss: 0.118
[epoch: 6, i: 6599] avg mini-batch loss: 0.174
[epoch: 6, i: 6699] avg mini-batch loss: 0.155
[epoch: 6, i: 6799] avg mini-batch loss: 0.135
[epoch: 6, i: 6899] avg mini-batch loss: 0.126
[epoch: 6, i: 6999] avg mini-batch loss: 0.144
[epoch: 6, i: 7099] avg mini-batch loss: 0.165
[epoch: 6, i: 7199] avg mini-batch loss: 0.113
[epoch: 6, i: 7299] avg mini-batch loss: 0.097
[epoch: 6, i: 7399] avg mini-batch loss: 0.134
[epoch: 6, i: 7499] avg mini-batch loss: 0.178
[epoch: 6, i: 7599] avg mini-batch loss: 0.123
[epoch: 6, i: 7699] avg mini-batch loss: 0.176
[epoch: 6, i: 7799] avg mini-batch loss: 0.153
[epoch: 6, i: 7899] avg mini-batch loss: 0.106

[epoch: 6, i: 7999] avg mini-batch loss: 0.111
[epoch: 6, i: 8099] avg mini-batch loss: 0.121
[epoch: 6, i: 8199] avg mini-batch loss: 0.118
[epoch: 6, i: 8299] avg mini-batch loss: 0.177
[epoch: 6, i: 8399] avg mini-batch loss: 0.126
[epoch: 6, i: 8499] avg mini-batch loss: 0.132
[epoch: 6, i: 8599] avg mini-batch loss: 0.127
[epoch: 6, i: 8699] avg mini-batch loss: 0.198
[epoch: 6, i: 8799] avg mini-batch loss: 0.168
[epoch: 6, i: 8899] avg mini-batch loss: 0.149
[epoch: 6, i: 8999] avg mini-batch loss: 0.169
[epoch: 6, i: 9099] avg mini-batch loss: 0.149
[epoch: 6, i: 9199] avg mini-batch loss: 0.220
[epoch: 6, i: 9299] avg mini-batch loss: 0.180
[epoch: 6, i: 9399] avg mini-batch loss: 0.201
[epoch: 6, i: 9499] avg mini-batch loss: 0.122
[epoch: 6, i: 9599] avg mini-batch loss: 0.115
[epoch: 6, i: 9699] avg mini-batch loss: 0.156
[epoch: 6, i: 9799] avg mini-batch loss: 0.119
[epoch: 6, i: 9899] avg mini-batch loss: 0.093
[epoch: 6, i: 9999] avg mini-batch loss: 0.116
[epoch: 6, i: 10099] avg mini-batch loss: 0.172
[epoch: 6, i: 10199] avg mini-batch loss: 0.141
[epoch: 6, i: 10299] avg mini-batch loss: 0.146
[epoch: 6, i: 10399] avg mini-batch loss: 0.158
[epoch: 6, i: 10499] avg mini-batch loss: 0.148
[epoch: 6, i: 10599] avg mini-batch loss: 0.171
[epoch: 6, i: 10699] avg mini-batch loss: 0.210
[epoch: 6, i: 10799] avg mini-batch loss: 0.178
[epoch: 6, i: 10899] avg mini-batch loss: 0.149
[epoch: 6, i: 10999] avg mini-batch loss: 0.168
[epoch: 6, i: 11099] avg mini-batch loss: 0.115
[epoch: 6, i: 11199] avg mini-batch loss: 0.138
[epoch: 6, i: 11299] avg mini-batch loss: 0.141
[epoch: 6, i: 11399] avg mini-batch loss: 0.165
[epoch: 6, i: 11499] avg mini-batch loss: 0.189
[epoch: 6, i: 11599] avg mini-batch loss: 0.184
[epoch: 6, i: 11699] avg mini-batch loss: 0.169
[epoch: 6, i: 11799] avg mini-batch loss: 0.185
[epoch: 6, i: 11899] avg mini-batch loss: 0.157
[epoch: 6, i: 11999] avg mini-batch loss: 0.180
[epoch: 6, i: 12099] avg mini-batch loss: 0.149
[epoch: 6, i: 12199] avg mini-batch loss: 0.110
[epoch: 6, i: 12299] avg mini-batch loss: 0.160
[epoch: 6, i: 12399] avg mini-batch loss: 0.169
[epoch: 6, i: 12499] avg mini-batch loss: 0.160
[epoch: 7, i: 99] avg mini-batch loss: 0.059
[epoch: 7, i: 199] avg mini-batch loss: 0.102

[epoch: 7, i: 299] avg mini-batch loss: 0.059
[epoch: 7, i: 399] avg mini-batch loss: 0.081
[epoch: 7, i: 499] avg mini-batch loss: 0.093
[epoch: 7, i: 599] avg mini-batch loss: 0.092
[epoch: 7, i: 699] avg mini-batch loss: 0.111
[epoch: 7, i: 799] avg mini-batch loss: 0.044
[epoch: 7, i: 899] avg mini-batch loss: 0.069
[epoch: 7, i: 999] avg mini-batch loss: 0.084
[epoch: 7, i: 1099] avg mini-batch loss: 0.082
[epoch: 7, i: 1199] avg mini-batch loss: 0.105
[epoch: 7, i: 1299] avg mini-batch loss: 0.064
[epoch: 7, i: 1399] avg mini-batch loss: 0.069
[epoch: 7, i: 1499] avg mini-batch loss: 0.060
[epoch: 7, i: 1599] avg mini-batch loss: 0.054
[epoch: 7, i: 1699] avg mini-batch loss: 0.076
[epoch: 7, i: 1799] avg mini-batch loss: 0.093
[epoch: 7, i: 1899] avg mini-batch loss: 0.056
[epoch: 7, i: 1999] avg mini-batch loss: 0.073
[epoch: 7, i: 2099] avg mini-batch loss: 0.070
[epoch: 7, i: 2199] avg mini-batch loss: 0.055
[epoch: 7, i: 2299] avg mini-batch loss: 0.134
[epoch: 7, i: 2399] avg mini-batch loss: 0.106
[epoch: 7, i: 2499] avg mini-batch loss: 0.115
[epoch: 7, i: 2599] avg mini-batch loss: 0.121
[epoch: 7, i: 2699] avg mini-batch loss: 0.092
[epoch: 7, i: 2799] avg mini-batch loss: 0.100
[epoch: 7, i: 2899] avg mini-batch loss: 0.104
[epoch: 7, i: 2999] avg mini-batch loss: 0.155
[epoch: 7, i: 3099] avg mini-batch loss: 0.119
[epoch: 7, i: 3199] avg mini-batch loss: 0.071
[epoch: 7, i: 3299] avg mini-batch loss: 0.030
[epoch: 7, i: 3399] avg mini-batch loss: 0.052
[epoch: 7, i: 3499] avg mini-batch loss: 0.047
[epoch: 7, i: 3599] avg mini-batch loss: 0.098
[epoch: 7, i: 3699] avg mini-batch loss: 0.113
[epoch: 7, i: 3799] avg mini-batch loss: 0.088
[epoch: 7, i: 3899] avg mini-batch loss: 0.068
[epoch: 7, i: 3999] avg mini-batch loss: 0.094
[epoch: 7, i: 4099] avg mini-batch loss: 0.089
[epoch: 7, i: 4199] avg mini-batch loss: 0.083
[epoch: 7, i: 4299] avg mini-batch loss: 0.068
[epoch: 7, i: 4399] avg mini-batch loss: 0.073
[epoch: 7, i: 4499] avg mini-batch loss: 0.065
[epoch: 7, i: 4599] avg mini-batch loss: 0.087
[epoch: 7, i: 4699] avg mini-batch loss: 0.065
[epoch: 7, i: 4799] avg mini-batch loss: 0.058
[epoch: 7, i: 4899] avg mini-batch loss: 0.047
[epoch: 7, i: 4999] avg mini-batch loss: 0.048

[epoch: 7, i: 5099] avg mini-batch loss: 0.120
[epoch: 7, i: 5199] avg mini-batch loss: 0.070
[epoch: 7, i: 5299] avg mini-batch loss: 0.073
[epoch: 7, i: 5399] avg mini-batch loss: 0.074
[epoch: 7, i: 5499] avg mini-batch loss: 0.082
[epoch: 7, i: 5599] avg mini-batch loss: 0.083
[epoch: 7, i: 5699] avg mini-batch loss: 0.080
[epoch: 7, i: 5799] avg mini-batch loss: 0.073
[epoch: 7, i: 5899] avg mini-batch loss: 0.131
[epoch: 7, i: 5999] avg mini-batch loss: 0.113
[epoch: 7, i: 6099] avg mini-batch loss: 0.099
[epoch: 7, i: 6199] avg mini-batch loss: 0.116
[epoch: 7, i: 6299] avg mini-batch loss: 0.082
[epoch: 7, i: 6399] avg mini-batch loss: 0.066
[epoch: 7, i: 6499] avg mini-batch loss: 0.061
[epoch: 7, i: 6599] avg mini-batch loss: 0.069
[epoch: 7, i: 6699] avg mini-batch loss: 0.090
[epoch: 7, i: 6799] avg mini-batch loss: 0.072
[epoch: 7, i: 6899] avg mini-batch loss: 0.088
[epoch: 7, i: 6999] avg mini-batch loss: 0.102
[epoch: 7, i: 7099] avg mini-batch loss: 0.107
[epoch: 7, i: 7199] avg mini-batch loss: 0.104
[epoch: 7, i: 7299] avg mini-batch loss: 0.087
[epoch: 7, i: 7399] avg mini-batch loss: 0.119
[epoch: 7, i: 7499] avg mini-batch loss: 0.105
[epoch: 7, i: 7599] avg mini-batch loss: 0.122
[epoch: 7, i: 7699] avg mini-batch loss: 0.098
[epoch: 7, i: 7799] avg mini-batch loss: 0.056
[epoch: 7, i: 7899] avg mini-batch loss: 0.092
[epoch: 7, i: 7999] avg mini-batch loss: 0.103
[epoch: 7, i: 8099] avg mini-batch loss: 0.106
[epoch: 7, i: 8199] avg mini-batch loss: 0.142
[epoch: 7, i: 8299] avg mini-batch loss: 0.126
[epoch: 7, i: 8399] avg mini-batch loss: 0.119
[epoch: 7, i: 8499] avg mini-batch loss: 0.103
[epoch: 7, i: 8599] avg mini-batch loss: 0.082
[epoch: 7, i: 8699] avg mini-batch loss: 0.060
[epoch: 7, i: 8799] avg mini-batch loss: 0.109
[epoch: 7, i: 8899] avg mini-batch loss: 0.065
[epoch: 7, i: 8999] avg mini-batch loss: 0.124
[epoch: 7, i: 9099] avg mini-batch loss: 0.092
[epoch: 7, i: 9199] avg mini-batch loss: 0.125
[epoch: 7, i: 9299] avg mini-batch loss: 0.140
[epoch: 7, i: 9399] avg mini-batch loss: 0.166
[epoch: 7, i: 9499] avg mini-batch loss: 0.148
[epoch: 7, i: 9599] avg mini-batch loss: 0.149
[epoch: 7, i: 9699] avg mini-batch loss: 0.165
[epoch: 7, i: 9799] avg mini-batch loss: 0.130

[epoch: 7, i: 9899] avg mini-batch loss: 0.117
[epoch: 7, i: 9999] avg mini-batch loss: 0.099
[epoch: 7, i: 10099] avg mini-batch loss: 0.108
[epoch: 7, i: 10199] avg mini-batch loss: 0.113
[epoch: 7, i: 10299] avg mini-batch loss: 0.121
[epoch: 7, i: 10399] avg mini-batch loss: 0.093
[epoch: 7, i: 10499] avg mini-batch loss: 0.124
[epoch: 7, i: 10599] avg mini-batch loss: 0.119
[epoch: 7, i: 10699] avg mini-batch loss: 0.171
[epoch: 7, i: 10799] avg mini-batch loss: 0.155
[epoch: 7, i: 10899] avg mini-batch loss: 0.178
[epoch: 7, i: 10999] avg mini-batch loss: 0.139
[epoch: 7, i: 11099] avg mini-batch loss: 0.127
[epoch: 7, i: 11199] avg mini-batch loss: 0.132
[epoch: 7, i: 11299] avg mini-batch loss: 0.119
[epoch: 7, i: 11399] avg mini-batch loss: 0.109
[epoch: 7, i: 11499] avg mini-batch loss: 0.110
[epoch: 7, i: 11599] avg mini-batch loss: 0.131
[epoch: 7, i: 11699] avg mini-batch loss: 0.128
[epoch: 7, i: 11799] avg mini-batch loss: 0.117
[epoch: 7, i: 11899] avg mini-batch loss: 0.131
[epoch: 7, i: 11999] avg mini-batch loss: 0.080
[epoch: 7, i: 12099] avg mini-batch loss: 0.103
[epoch: 7, i: 12199] avg mini-batch loss: 0.139
[epoch: 7, i: 12299] avg mini-batch loss: 0.139
[epoch: 7, i: 12399] avg mini-batch loss: 0.070
[epoch: 7, i: 12499] avg mini-batch loss: 0.090
[epoch: 8, i: 99] avg mini-batch loss: 0.047
[epoch: 8, i: 199] avg mini-batch loss: 0.054
[epoch: 8, i: 299] avg mini-batch loss: 0.056
[epoch: 8, i: 399] avg mini-batch loss: 0.053
[epoch: 8, i: 499] avg mini-batch loss: 0.077
[epoch: 8, i: 599] avg mini-batch loss: 0.029
[epoch: 8, i: 699] avg mini-batch loss: 0.045
[epoch: 8, i: 799] avg mini-batch loss: 0.069
[epoch: 8, i: 899] avg mini-batch loss: 0.047
[epoch: 8, i: 999] avg mini-batch loss: 0.052
[epoch: 8, i: 1099] avg mini-batch loss: 0.052
[epoch: 8, i: 1199] avg mini-batch loss: 0.029
[epoch: 8, i: 1299] avg mini-batch loss: 0.046
[epoch: 8, i: 1399] avg mini-batch loss: 0.026
[epoch: 8, i: 1499] avg mini-batch loss: 0.028
[epoch: 8, i: 1599] avg mini-batch loss: 0.025
[epoch: 8, i: 1699] avg mini-batch loss: 0.061
[epoch: 8, i: 1799] avg mini-batch loss: 0.043
[epoch: 8, i: 1899] avg mini-batch loss: 0.050
[epoch: 8, i: 1999] avg mini-batch loss: 0.035
[epoch: 8, i: 2099] avg mini-batch loss: 0.026

[epoch: 8, i: 2199] avg mini-batch loss: 0.034
[epoch: 8, i: 2299] avg mini-batch loss: 0.040
[epoch: 8, i: 2399] avg mini-batch loss: 0.067
[epoch: 8, i: 2499] avg mini-batch loss: 0.034
[epoch: 8, i: 2599] avg mini-batch loss: 0.037
[epoch: 8, i: 2699] avg mini-batch loss: 0.034
[epoch: 8, i: 2799] avg mini-batch loss: 0.039
[epoch: 8, i: 2899] avg mini-batch loss: 0.042
[epoch: 8, i: 2999] avg mini-batch loss: 0.030
[epoch: 8, i: 3099] avg mini-batch loss: 0.046
[epoch: 8, i: 3199] avg mini-batch loss: 0.042
[epoch: 8, i: 3299] avg mini-batch loss: 0.066
[epoch: 8, i: 3399] avg mini-batch loss: 0.066
[epoch: 8, i: 3499] avg mini-batch loss: 0.087
[epoch: 8, i: 3599] avg mini-batch loss: 0.108
[epoch: 8, i: 3699] avg mini-batch loss: 0.061
[epoch: 8, i: 3799] avg mini-batch loss: 0.050
[epoch: 8, i: 3899] avg mini-batch loss: 0.063
[epoch: 8, i: 3999] avg mini-batch loss: 0.074
[epoch: 8, i: 4099] avg mini-batch loss: 0.085
[epoch: 8, i: 4199] avg mini-batch loss: 0.110
[epoch: 8, i: 4299] avg mini-batch loss: 0.127
[epoch: 8, i: 4399] avg mini-batch loss: 0.082
[epoch: 8, i: 4499] avg mini-batch loss: 0.053
[epoch: 8, i: 4599] avg mini-batch loss: 0.051
[epoch: 8, i: 4699] avg mini-batch loss: 0.044
[epoch: 8, i: 4799] avg mini-batch loss: 0.031
[epoch: 8, i: 4899] avg mini-batch loss: 0.044
[epoch: 8, i: 4999] avg mini-batch loss: 0.081
[epoch: 8, i: 5099] avg mini-batch loss: 0.090
[epoch: 8, i: 5199] avg mini-batch loss: 0.071
[epoch: 8, i: 5299] avg mini-batch loss: 0.111
[epoch: 8, i: 5399] avg mini-batch loss: 0.036
[epoch: 8, i: 5499] avg mini-batch loss: 0.041
[epoch: 8, i: 5599] avg mini-batch loss: 0.031
[epoch: 8, i: 5699] avg mini-batch loss: 0.063
[epoch: 8, i: 5799] avg mini-batch loss: 0.055
[epoch: 8, i: 5899] avg mini-batch loss: 0.077
[epoch: 8, i: 5999] avg mini-batch loss: 0.099
[epoch: 8, i: 6099] avg mini-batch loss: 0.074
[epoch: 8, i: 6199] avg mini-batch loss: 0.046
[epoch: 8, i: 6299] avg mini-batch loss: 0.055
[epoch: 8, i: 6399] avg mini-batch loss: 0.101
[epoch: 8, i: 6499] avg mini-batch loss: 0.133
[epoch: 8, i: 6599] avg mini-batch loss: 0.084
[epoch: 8, i: 6699] avg mini-batch loss: 0.086
[epoch: 8, i: 6799] avg mini-batch loss: 0.083
[epoch: 8, i: 6899] avg mini-batch loss: 0.102

[epoch: 8, i: 6999] avg mini-batch loss: 0.061
[epoch: 8, i: 7099] avg mini-batch loss: 0.059
[epoch: 8, i: 7199] avg mini-batch loss: 0.052
[epoch: 8, i: 7299] avg mini-batch loss: 0.070
[epoch: 8, i: 7399] avg mini-batch loss: 0.084
[epoch: 8, i: 7499] avg mini-batch loss: 0.053
[epoch: 8, i: 7599] avg mini-batch loss: 0.044
[epoch: 8, i: 7699] avg mini-batch loss: 0.048
[epoch: 8, i: 7799] avg mini-batch loss: 0.031
[epoch: 8, i: 7899] avg mini-batch loss: 0.079
[epoch: 8, i: 7999] avg mini-batch loss: 0.046
[epoch: 8, i: 8099] avg mini-batch loss: 0.063
[epoch: 8, i: 8199] avg mini-batch loss: 0.098
[epoch: 8, i: 8299] avg mini-batch loss: 0.060
[epoch: 8, i: 8399] avg mini-batch loss: 0.073
[epoch: 8, i: 8499] avg mini-batch loss: 0.067
[epoch: 8, i: 8599] avg mini-batch loss: 0.104
[epoch: 8, i: 8699] avg mini-batch loss: 0.117
[epoch: 8, i: 8799] avg mini-batch loss: 0.081
[epoch: 8, i: 8899] avg mini-batch loss: 0.062
[epoch: 8, i: 8999] avg mini-batch loss: 0.067
[epoch: 8, i: 9099] avg mini-batch loss: 0.058
[epoch: 8, i: 9199] avg mini-batch loss: 0.094
[epoch: 8, i: 9299] avg mini-batch loss: 0.084
[epoch: 8, i: 9399] avg mini-batch loss: 0.137
[epoch: 8, i: 9499] avg mini-batch loss: 0.114
[epoch: 8, i: 9599] avg mini-batch loss: 0.115
[epoch: 8, i: 9699] avg mini-batch loss: 0.112
[epoch: 8, i: 9799] avg mini-batch loss: 0.070
[epoch: 8, i: 9899] avg mini-batch loss: 0.127
[epoch: 8, i: 9999] avg mini-batch loss: 0.070
[epoch: 8, i: 10099] avg mini-batch loss: 0.104
[epoch: 8, i: 10199] avg mini-batch loss: 0.083
[epoch: 8, i: 10299] avg mini-batch loss: 0.078
[epoch: 8, i: 10399] avg mini-batch loss: 0.087
[epoch: 8, i: 10499] avg mini-batch loss: 0.049
[epoch: 8, i: 10599] avg mini-batch loss: 0.053
[epoch: 8, i: 10699] avg mini-batch loss: 0.071
[epoch: 8, i: 10799] avg mini-batch loss: 0.045
[epoch: 8, i: 10899] avg mini-batch loss: 0.054
[epoch: 8, i: 10999] avg mini-batch loss: 0.083
[epoch: 8, i: 11099] avg mini-batch loss: 0.040
[epoch: 8, i: 11199] avg mini-batch loss: 0.061
[epoch: 8, i: 11299] avg mini-batch loss: 0.118
[epoch: 8, i: 11399] avg mini-batch loss: 0.152
[epoch: 8, i: 11499] avg mini-batch loss: 0.113
[epoch: 8, i: 11599] avg mini-batch loss: 0.101
[epoch: 8, i: 11699] avg mini-batch loss: 0.084

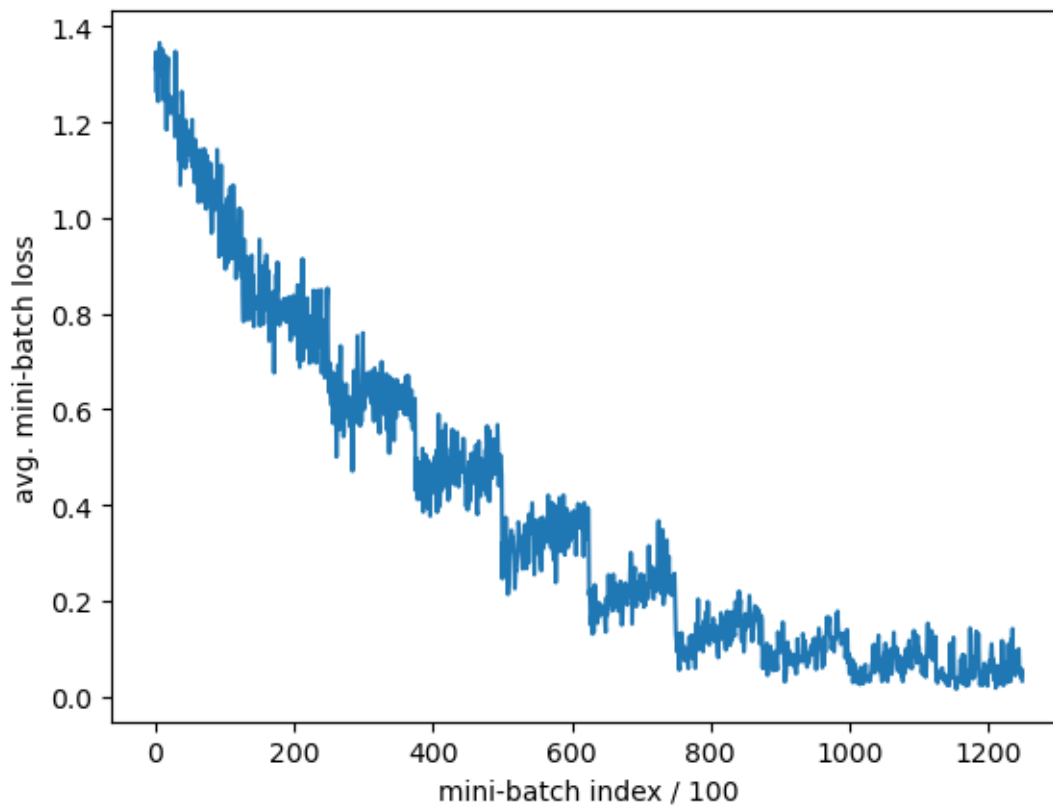
[epoch: 8, i: 11799] avg mini-batch loss: 0.130
[epoch: 8, i: 11899] avg mini-batch loss: 0.137
[epoch: 8, i: 11999] avg mini-batch loss: 0.078
[epoch: 8, i: 12099] avg mini-batch loss: 0.076
[epoch: 8, i: 12199] avg mini-batch loss: 0.093
[epoch: 8, i: 12299] avg mini-batch loss: 0.093
[epoch: 8, i: 12399] avg mini-batch loss: 0.083
[epoch: 8, i: 12499] avg mini-batch loss: 0.125
[epoch: 9, i: 99] avg mini-batch loss: 0.034
[epoch: 9, i: 199] avg mini-batch loss: 0.032
[epoch: 9, i: 299] avg mini-batch loss: 0.028
[epoch: 9, i: 399] avg mini-batch loss: 0.044
[epoch: 9, i: 499] avg mini-batch loss: 0.052
[epoch: 9, i: 599] avg mini-batch loss: 0.049
[epoch: 9, i: 699] avg mini-batch loss: 0.040
[epoch: 9, i: 799] avg mini-batch loss: 0.057
[epoch: 9, i: 899] avg mini-batch loss: 0.058
[epoch: 9, i: 999] avg mini-batch loss: 0.059
[epoch: 9, i: 1099] avg mini-batch loss: 0.034
[epoch: 9, i: 1199] avg mini-batch loss: 0.046
[epoch: 9, i: 1299] avg mini-batch loss: 0.032
[epoch: 9, i: 1399] avg mini-batch loss: 0.045
[epoch: 9, i: 1499] avg mini-batch loss: 0.029
[epoch: 9, i: 1599] avg mini-batch loss: 0.034
[epoch: 9, i: 1699] avg mini-batch loss: 0.023
[epoch: 9, i: 1799] avg mini-batch loss: 0.032
[epoch: 9, i: 1899] avg mini-batch loss: 0.044
[epoch: 9, i: 1999] avg mini-batch loss: 0.023
[epoch: 9, i: 2099] avg mini-batch loss: 0.111
[epoch: 9, i: 2199] avg mini-batch loss: 0.050
[epoch: 9, i: 2299] avg mini-batch loss: 0.058
[epoch: 9, i: 2399] avg mini-batch loss: 0.038
[epoch: 9, i: 2499] avg mini-batch loss: 0.069
[epoch: 9, i: 2599] avg mini-batch loss: 0.123
[epoch: 9, i: 2699] avg mini-batch loss: 0.043
[epoch: 9, i: 2799] avg mini-batch loss: 0.040
[epoch: 9, i: 2899] avg mini-batch loss: 0.035
[epoch: 9, i: 2999] avg mini-batch loss: 0.015
[epoch: 9, i: 3099] avg mini-batch loss: 0.034
[epoch: 9, i: 3199] avg mini-batch loss: 0.040
[epoch: 9, i: 3299] avg mini-batch loss: 0.033
[epoch: 9, i: 3399] avg mini-batch loss: 0.034
[epoch: 9, i: 3499] avg mini-batch loss: 0.033
[epoch: 9, i: 3599] avg mini-batch loss: 0.076
[epoch: 9, i: 3699] avg mini-batch loss: 0.071
[epoch: 9, i: 3799] avg mini-batch loss: 0.027
[epoch: 9, i: 3899] avg mini-batch loss: 0.087
[epoch: 9, i: 3999] avg mini-batch loss: 0.041

[epoch: 9, i: 4099] avg mini-batch loss: 0.087
[epoch: 9, i: 4199] avg mini-batch loss: 0.056
[epoch: 9, i: 4299] avg mini-batch loss: 0.032
[epoch: 9, i: 4399] avg mini-batch loss: 0.022
[epoch: 9, i: 4499] avg mini-batch loss: 0.038
[epoch: 9, i: 4599] avg mini-batch loss: 0.053
[epoch: 9, i: 4699] avg mini-batch loss: 0.048
[epoch: 9, i: 4799] avg mini-batch loss: 0.024
[epoch: 9, i: 4899] avg mini-batch loss: 0.068
[epoch: 9, i: 4999] avg mini-batch loss: 0.142
[epoch: 9, i: 5099] avg mini-batch loss: 0.068
[epoch: 9, i: 5199] avg mini-batch loss: 0.048
[epoch: 9, i: 5299] avg mini-batch loss: 0.039
[epoch: 9, i: 5399] avg mini-batch loss: 0.037
[epoch: 9, i: 5499] avg mini-batch loss: 0.081
[epoch: 9, i: 5599] avg mini-batch loss: 0.040
[epoch: 9, i: 5699] avg mini-batch loss: 0.042
[epoch: 9, i: 5799] avg mini-batch loss: 0.128
[epoch: 9, i: 5899] avg mini-batch loss: 0.136
[epoch: 9, i: 5999] avg mini-batch loss: 0.060
[epoch: 9, i: 6099] avg mini-batch loss: 0.080
[epoch: 9, i: 6199] avg mini-batch loss: 0.131
[epoch: 9, i: 6299] avg mini-batch loss: 0.066
[epoch: 9, i: 6399] avg mini-batch loss: 0.024
[epoch: 9, i: 6499] avg mini-batch loss: 0.049
[epoch: 9, i: 6599] avg mini-batch loss: 0.038
[epoch: 9, i: 6699] avg mini-batch loss: 0.048
[epoch: 9, i: 6799] avg mini-batch loss: 0.023
[epoch: 9, i: 6899] avg mini-batch loss: 0.047
[epoch: 9, i: 6999] avg mini-batch loss: 0.069
[epoch: 9, i: 7099] avg mini-batch loss: 0.046
[epoch: 9, i: 7199] avg mini-batch loss: 0.062
[epoch: 9, i: 7299] avg mini-batch loss: 0.035
[epoch: 9, i: 7399] avg mini-batch loss: 0.058
[epoch: 9, i: 7499] avg mini-batch loss: 0.022
[epoch: 9, i: 7599] avg mini-batch loss: 0.064
[epoch: 9, i: 7699] avg mini-batch loss: 0.066
[epoch: 9, i: 7799] avg mini-batch loss: 0.074
[epoch: 9, i: 7899] avg mini-batch loss: 0.052
[epoch: 9, i: 7999] avg mini-batch loss: 0.091
[epoch: 9, i: 8099] avg mini-batch loss: 0.109
[epoch: 9, i: 8199] avg mini-batch loss: 0.070
[epoch: 9, i: 8299] avg mini-batch loss: 0.078
[epoch: 9, i: 8399] avg mini-batch loss: 0.034
[epoch: 9, i: 8499] avg mini-batch loss: 0.047
[epoch: 9, i: 8599] avg mini-batch loss: 0.046
[epoch: 9, i: 8699] avg mini-batch loss: 0.017
[epoch: 9, i: 8799] avg mini-batch loss: 0.060

```
[epoch: 9, i: 8899] avg mini-batch loss: 0.065
[epoch: 9, i: 8999] avg mini-batch loss: 0.050
[epoch: 9, i: 9099] avg mini-batch loss: 0.025
[epoch: 9, i: 9199] avg mini-batch loss: 0.026
[epoch: 9, i: 9299] avg mini-batch loss: 0.044
[epoch: 9, i: 9399] avg mini-batch loss: 0.056
[epoch: 9, i: 9499] avg mini-batch loss: 0.065
[epoch: 9, i: 9599] avg mini-batch loss: 0.116
[epoch: 9, i: 9699] avg mini-batch loss: 0.077
[epoch: 9, i: 9799] avg mini-batch loss: 0.022
[epoch: 9, i: 9899] avg mini-batch loss: 0.041
[epoch: 9, i: 9999] avg mini-batch loss: 0.051
[epoch: 9, i: 10099] avg mini-batch loss: 0.068
[epoch: 9, i: 10199] avg mini-batch loss: 0.041
[epoch: 9, i: 10299] avg mini-batch loss: 0.029
[epoch: 9, i: 10399] avg mini-batch loss: 0.091
[epoch: 9, i: 10499] avg mini-batch loss: 0.109
[epoch: 9, i: 10599] avg mini-batch loss: 0.091
[epoch: 9, i: 10699] avg mini-batch loss: 0.086
[epoch: 9, i: 10799] avg mini-batch loss: 0.031
[epoch: 9, i: 10899] avg mini-batch loss: 0.118
[epoch: 9, i: 10999] avg mini-batch loss: 0.078
[epoch: 9, i: 11099] avg mini-batch loss: 0.141
[epoch: 9, i: 11199] avg mini-batch loss: 0.057
[epoch: 9, i: 11299] avg mini-batch loss: 0.097
[epoch: 9, i: 11399] avg mini-batch loss: 0.094
[epoch: 9, i: 11499] avg mini-batch loss: 0.043
[epoch: 9, i: 11599] avg mini-batch loss: 0.084
[epoch: 9, i: 11699] avg mini-batch loss: 0.053
[epoch: 9, i: 11799] avg mini-batch loss: 0.079
[epoch: 9, i: 11899] avg mini-batch loss: 0.073
[epoch: 9, i: 11999] avg mini-batch loss: 0.099
[epoch: 9, i: 12099] avg mini-batch loss: 0.051
[epoch: 9, i: 12199] avg mini-batch loss: 0.049
[epoch: 9, i: 12299] avg mini-batch loss: 0.039
[epoch: 9, i: 12399] avg mini-batch loss: 0.057
[epoch: 9, i: 12499] avg mini-batch loss: 0.032
Finished Training.
```

Training Loss Curve

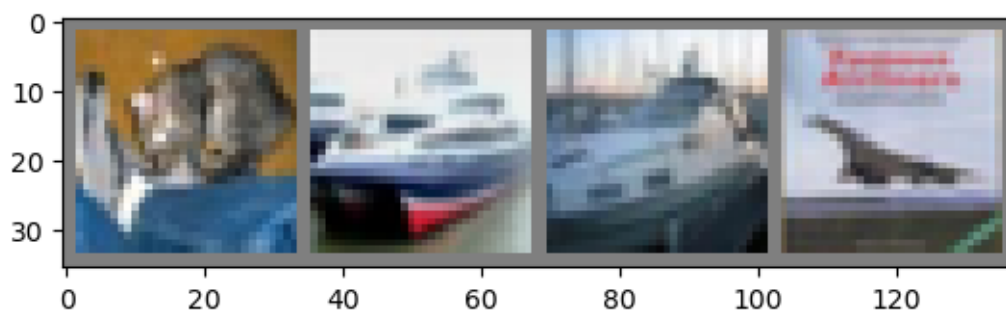
```
[11]: plt.plot(avg_losses)
plt.xlabel('mini-batch index / {}'.format(print_freq))
plt.ylabel('avg. mini-batch loss')
plt.show()
```



Evaluate on Test Dataset

```
[12]: # Check several images.
dataiter = iter(testloader)
images, labels = next(dataiter)
imshow(torchvision.utils.make_grid(images))
print('GroundTruth: ', ' '.join('%5s' % classes[labels[j]] for j in range(4)))
outputs = net(images.to(device))
_, predicted = torch.max(outputs, 1)

print('Predicted: ', ' '.join('%5s' % classes[predicted[j]]
                               for j in range(4)))
```



GroundTruth: cat ship ship plane
Predicted: cat ship ship plane

```
[13]: # Get test accuracy.
correct = 0
total = 0
with torch.no_grad():
    for data in testloader:
        images, labels = data
        images, labels = images.to(device), labels.to(device)
        outputs = net(images)
        _, predicted = torch.max(outputs.data, 1)
        total += labels.size(0)
        correct += (predicted == labels).sum().item()

print('Accuracy of the network on the 10000 test images: %d %%' % (
    100 * correct / total))
```

Accuracy of the network on the 10000 test images: 78 %

```
[14]: # Get test accuracy for each class.
class_correct = list(0. for i in range(10))
class_total = list(0. for i in range(10))
with torch.no_grad():
    for data in testloader:
        images, labels = data
        images, labels = images.to(device), labels.to(device)
        outputs = net(images)
        _, predicted = torch.max(outputs, 1)
        c = (predicted == labels).squeeze()
        for i in range(4):
            label = labels[i]
            class_correct[label] += c[i].item()
            class_total[label] += 1

for i in range(10):
    print('Accuracy of %5s : %2d %%' % (
        classes[i], 100 * class_correct[i] / class_total[i]))
```

Accuracy of plane : 78 %
Accuracy of car : 88 %
Accuracy of bird : 72 %
Accuracy of cat : 53 %
Accuracy of deer : 75 %
Accuracy of dog : 70 %
Accuracy of frog : 87 %

Accuracy of horse : 84 %
Accuracy of ship : 87 %
Accuracy of truck : 88 %

CNN # 2 Using TensorFlow

- Train dataset on a new CNN model compile with Adamax optimizer

```
[21]: from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Dense, Flatten,
↳Dropout, BatchNormalization, LeakyReLU

model = Sequential()
model.add(Conv2D(16, (3,3), strides=(1,1), padding='same', use_bias=False,
↳input_shape=(32,32,3)))
model.add(BatchNormalization())
model.add(LeakyReLU(negative_slope=0.1))
model.add(MaxPooling2D(pool_size=(2, 2),padding='same'))
model.add(Dropout(0.25))

for i in range(0,3):
    model.add(Conv2D(32*(2**i), (3,3), strides=(1,1), padding='same',
↳use_bias=False))
    model.add(BatchNormalization())
    model.add(LeakyReLU(negative_slope=0.1))
    model.add(MaxPooling2D(pool_size=(2, 2),padding='same'))
    model.add(Dropout(0.25))

model.add(Conv2D(512, (3,3), strides=(1,1), padding='same', use_bias=False))
model.add(BatchNormalization())
model.add(LeakyReLU(negative_slope=0.1))
model.add(MaxPooling2D(pool_size=(2, 2), strides=(1,1), padding='same'))
model.add(Dropout(0.25))

model.add(Flatten())
model.add(Dense(10, activation='softmax'))

# Compile the model
model.compile(loss='categorical_crossentropy',
              optimizer='adamax',
              metrics=['categorical_accuracy'])
model.summary()
```

Model: "sequential_5"

Layer (type)

Output Shape

Param #

conv2d_12 (Conv2D)	(None, 32, 32, 16)	432
batch_normalization_11 (BatchNormalization)	(None, 32, 32, 16)	64
leaky_re_lu_10 (LeakyReLU)	(None, 32, 32, 16)	0
max_pooling2d_10 (MaxPooling2D)	(None, 16, 16, 16)	0
dropout_10 (Dropout)	(None, 16, 16, 16)	0
conv2d_13 (Conv2D)	(None, 16, 16, 32)	4,608
batch_normalization_12 (BatchNormalization)	(None, 16, 16, 32)	128
leaky_re_lu_11 (LeakyReLU)	(None, 16, 16, 32)	0
max_pooling2d_11 (MaxPooling2D)	(None, 8, 8, 32)	0
dropout_11 (Dropout)	(None, 8, 8, 32)	0
conv2d_14 (Conv2D)	(None, 8, 8, 64)	18,432
batch_normalization_13 (BatchNormalization)	(None, 8, 8, 64)	256
leaky_re_lu_12 (LeakyReLU)	(None, 8, 8, 64)	0
max_pooling2d_12 (MaxPooling2D)	(None, 4, 4, 64)	0
dropout_12 (Dropout)	(None, 4, 4, 64)	0
conv2d_15 (Conv2D)	(None, 4, 4, 128)	73,728
batch_normalization_14 (BatchNormalization)	(None, 4, 4, 128)	512
leaky_re_lu_13 (LeakyReLU)	(None, 4, 4, 128)	0
max_pooling2d_13 (MaxPooling2D)	(None, 2, 2, 128)	0
dropout_13 (Dropout)	(None, 2, 2, 128)	0
conv2d_16 (Conv2D)	(None, 2, 2, 512)	589,824
batch_normalization_15 (BatchNormalization)	(None, 2, 2, 512)	2,048

leaky_re_lu_14 (LeakyReLU)	(None, 2, 2, 512)	0
max_pooling2d_14 (MaxPooling2D)	(None, 2, 2, 512)	0
dropout_14 (Dropout)	(None, 2, 2, 512)	0
flatten_2 (Flatten)	(None, 2048)	0
dense_2 (Dense)	(None, 10)	20,490

Total params: 710,522 (2.71 MB)

Trainable params: 709,018 (2.70 MB)

Non-trainable params: 1,504 (5.88 KB)

```
[22]: avg_losses = []      # Avg. losses.
epochs = 10             # Total epochs.
print_freq = 100        # Print frequency.

for epoch in range(epochs): # Loop over the dataset multiple times.
    running_loss = 0.0      # Initialize running loss.
    for i, data in enumerate(trainloader, 0):
        # Get the inputs.
        inputs, labels = data

        # Move the inputs to the specified device.
        inputs, labels = inputs.to(device), labels.to(device)

        # Zero the parameter gradients.
        opt.zero_grad()

        # Forward step.
        outputs = net(inputs)
        loss = loss_func(outputs, labels)

        # Backward step.
        loss.backward()

        # Optimization step (update the parameters).
        opt.step()

        # Print statistics.
```

```

        running_loss += loss.item()
        if i % print_freq == print_freq - 1: # Print every several mini-batches.
            avg_loss = running_loss / print_freq
            print('[epoch: {}], i: {:5d}] avg mini-batch loss: {:.3f}'.format(
                epoch, i, avg_loss))
            avg_losses.append(avg_loss)
            running_loss = 0.0

print('Finished Training.')

```

```

[epoch: 0, i:    99] avg mini-batch loss: 0.035
[epoch: 0, i:   199] avg mini-batch loss: 0.029
[epoch: 0, i:   299] avg mini-batch loss: 0.027
[epoch: 0, i:   399] avg mini-batch loss: 0.049
[epoch: 0, i:   499] avg mini-batch loss: 0.019
[epoch: 0, i:   599] avg mini-batch loss: 0.034
[epoch: 0, i:   699] avg mini-batch loss: 0.034
[epoch: 0, i:   799] avg mini-batch loss: 0.020
[epoch: 0, i:   899] avg mini-batch loss: 0.056
[epoch: 0, i:   999] avg mini-batch loss: 0.022
[epoch: 0, i:  1099] avg mini-batch loss: 0.018
[epoch: 0, i:  1199] avg mini-batch loss: 0.050
[epoch: 0, i:  1299] avg mini-batch loss: 0.041
[epoch: 0, i:  1399] avg mini-batch loss: 0.026
[epoch: 0, i:  1499] avg mini-batch loss: 0.034
[epoch: 0, i:  1599] avg mini-batch loss: 0.089
[epoch: 0, i:  1699] avg mini-batch loss: 0.054
[epoch: 0, i:  1799] avg mini-batch loss: 0.033
[epoch: 0, i:  1899] avg mini-batch loss: 0.108
[epoch: 0, i:  1999] avg mini-batch loss: 0.051
[epoch: 0, i:  2099] avg mini-batch loss: 0.063
[epoch: 0, i:  2199] avg mini-batch loss: 0.094
[epoch: 0, i:  2299] avg mini-batch loss: 0.070
[epoch: 0, i:  2399] avg mini-batch loss: 0.027
[epoch: 0, i:  2499] avg mini-batch loss: 0.021
[epoch: 0, i:  2599] avg mini-batch loss: 0.012
[epoch: 0, i:  2699] avg mini-batch loss: 0.013
[epoch: 0, i:  2799] avg mini-batch loss: 0.039
[epoch: 0, i:  2899] avg mini-batch loss: 0.024
[epoch: 0, i:  2999] avg mini-batch loss: 0.034
[epoch: 0, i:  3099] avg mini-batch loss: 0.051
[epoch: 0, i:  3199] avg mini-batch loss: 0.054
[epoch: 0, i:  3299] avg mini-batch loss: 0.035
[epoch: 0, i:  3399] avg mini-batch loss: 0.038
[epoch: 0, i:  3499] avg mini-batch loss: 0.029
[epoch: 0, i:  3599] avg mini-batch loss: 0.048
[epoch: 0, i:  3699] avg mini-batch loss: 0.056
[epoch: 0, i:  3799] avg mini-batch loss: 0.026

```

[epoch: 0, i: 3899] avg mini-batch loss: 0.015
[epoch: 0, i: 3999] avg mini-batch loss: 0.025
[epoch: 0, i: 4099] avg mini-batch loss: 0.020
[epoch: 0, i: 4199] avg mini-batch loss: 0.044
[epoch: 0, i: 4299] avg mini-batch loss: 0.054
[epoch: 0, i: 4399] avg mini-batch loss: 0.080
[epoch: 0, i: 4499] avg mini-batch loss: 0.047
[epoch: 0, i: 4599] avg mini-batch loss: 0.027
[epoch: 0, i: 4699] avg mini-batch loss: 0.077
[epoch: 0, i: 4799] avg mini-batch loss: 0.049
[epoch: 0, i: 4899] avg mini-batch loss: 0.045
[epoch: 0, i: 4999] avg mini-batch loss: 0.049
[epoch: 0, i: 5099] avg mini-batch loss: 0.067
[epoch: 0, i: 5199] avg mini-batch loss: 0.069
[epoch: 0, i: 5299] avg mini-batch loss: 0.056
[epoch: 0, i: 5399] avg mini-batch loss: 0.016
[epoch: 0, i: 5499] avg mini-batch loss: 0.034
[epoch: 0, i: 5599] avg mini-batch loss: 0.073
[epoch: 0, i: 5699] avg mini-batch loss: 0.052
[epoch: 0, i: 5799] avg mini-batch loss: 0.035
[epoch: 0, i: 5899] avg mini-batch loss: 0.045
[epoch: 0, i: 5999] avg mini-batch loss: 0.067
[epoch: 0, i: 6099] avg mini-batch loss: 0.048
[epoch: 0, i: 6199] avg mini-batch loss: 0.056
[epoch: 0, i: 6299] avg mini-batch loss: 0.041
[epoch: 0, i: 6399] avg mini-batch loss: 0.035
[epoch: 0, i: 6499] avg mini-batch loss: 0.069
[epoch: 0, i: 6599] avg mini-batch loss: 0.030
[epoch: 0, i: 6699] avg mini-batch loss: 0.042
[epoch: 0, i: 6799] avg mini-batch loss: 0.047
[epoch: 0, i: 6899] avg mini-batch loss: 0.021
[epoch: 0, i: 6999] avg mini-batch loss: 0.023
[epoch: 0, i: 7099] avg mini-batch loss: 0.010
[epoch: 0, i: 7199] avg mini-batch loss: 0.053
[epoch: 0, i: 7299] avg mini-batch loss: 0.055
[epoch: 0, i: 7399] avg mini-batch loss: 0.035
[epoch: 0, i: 7499] avg mini-batch loss: 0.075
[epoch: 0, i: 7599] avg mini-batch loss: 0.069
[epoch: 0, i: 7699] avg mini-batch loss: 0.070
[epoch: 0, i: 7799] avg mini-batch loss: 0.060
[epoch: 0, i: 7899] avg mini-batch loss: 0.089
[epoch: 0, i: 7999] avg mini-batch loss: 0.035
[epoch: 0, i: 8099] avg mini-batch loss: 0.076
[epoch: 0, i: 8199] avg mini-batch loss: 0.034
[epoch: 0, i: 8299] avg mini-batch loss: 0.037
[epoch: 0, i: 8399] avg mini-batch loss: 0.039
[epoch: 0, i: 8499] avg mini-batch loss: 0.039
[epoch: 0, i: 8599] avg mini-batch loss: 0.020

[epoch: 0, i: 8699] avg mini-batch loss: 0.040
[epoch: 0, i: 8799] avg mini-batch loss: 0.036
[epoch: 0, i: 8899] avg mini-batch loss: 0.049
[epoch: 0, i: 8999] avg mini-batch loss: 0.056
[epoch: 0, i: 9099] avg mini-batch loss: 0.040
[epoch: 0, i: 9199] avg mini-batch loss: 0.123
[epoch: 0, i: 9299] avg mini-batch loss: 0.089
[epoch: 0, i: 9399] avg mini-batch loss: 0.035
[epoch: 0, i: 9499] avg mini-batch loss: 0.058
[epoch: 0, i: 9599] avg mini-batch loss: 0.031
[epoch: 0, i: 9699] avg mini-batch loss: 0.071
[epoch: 0, i: 9799] avg mini-batch loss: 0.064
[epoch: 0, i: 9899] avg mini-batch loss: 0.065
[epoch: 0, i: 9999] avg mini-batch loss: 0.080
[epoch: 0, i: 10099] avg mini-batch loss: 0.068
[epoch: 0, i: 10199] avg mini-batch loss: 0.095
[epoch: 0, i: 10299] avg mini-batch loss: 0.032
[epoch: 0, i: 10399] avg mini-batch loss: 0.025
[epoch: 0, i: 10499] avg mini-batch loss: 0.041
[epoch: 0, i: 10599] avg mini-batch loss: 0.045
[epoch: 0, i: 10699] avg mini-batch loss: 0.080
[epoch: 0, i: 10799] avg mini-batch loss: 0.050
[epoch: 0, i: 10899] avg mini-batch loss: 0.089
[epoch: 0, i: 10999] avg mini-batch loss: 0.052
[epoch: 0, i: 11099] avg mini-batch loss: 0.054
[epoch: 0, i: 11199] avg mini-batch loss: 0.031
[epoch: 0, i: 11299] avg mini-batch loss: 0.058
[epoch: 0, i: 11399] avg mini-batch loss: 0.063
[epoch: 0, i: 11499] avg mini-batch loss: 0.048
[epoch: 0, i: 11599] avg mini-batch loss: 0.023
[epoch: 0, i: 11699] avg mini-batch loss: 0.062
[epoch: 0, i: 11799] avg mini-batch loss: 0.066
[epoch: 0, i: 11899] avg mini-batch loss: 0.057
[epoch: 0, i: 11999] avg mini-batch loss: 0.069
[epoch: 0, i: 12099] avg mini-batch loss: 0.046
[epoch: 0, i: 12199] avg mini-batch loss: 0.063
[epoch: 0, i: 12299] avg mini-batch loss: 0.059
[epoch: 0, i: 12399] avg mini-batch loss: 0.054
[epoch: 0, i: 12499] avg mini-batch loss: 0.051
[epoch: 1, i: 99] avg mini-batch loss: 0.049
[epoch: 1, i: 199] avg mini-batch loss: 0.038
[epoch: 1, i: 299] avg mini-batch loss: 0.024
[epoch: 1, i: 399] avg mini-batch loss: 0.043
[epoch: 1, i: 499] avg mini-batch loss: 0.025
[epoch: 1, i: 599] avg mini-batch loss: 0.030
[epoch: 1, i: 699] avg mini-batch loss: 0.027
[epoch: 1, i: 799] avg mini-batch loss: 0.026
[epoch: 1, i: 899] avg mini-batch loss: 0.015

[epoch: 1, i: 999] avg mini-batch loss: 0.048
[epoch: 1, i: 1099] avg mini-batch loss: 0.036
[epoch: 1, i: 1199] avg mini-batch loss: 0.015
[epoch: 1, i: 1299] avg mini-batch loss: 0.034
[epoch: 1, i: 1399] avg mini-batch loss: 0.019
[epoch: 1, i: 1499] avg mini-batch loss: 0.033
[epoch: 1, i: 1599] avg mini-batch loss: 0.025
[epoch: 1, i: 1699] avg mini-batch loss: 0.039
[epoch: 1, i: 1799] avg mini-batch loss: 0.039
[epoch: 1, i: 1899] avg mini-batch loss: 0.027
[epoch: 1, i: 1999] avg mini-batch loss: 0.055
[epoch: 1, i: 2099] avg mini-batch loss: 0.043
[epoch: 1, i: 2199] avg mini-batch loss: 0.033
[epoch: 1, i: 2299] avg mini-batch loss: 0.035
[epoch: 1, i: 2399] avg mini-batch loss: 0.016
[epoch: 1, i: 2499] avg mini-batch loss: 0.026
[epoch: 1, i: 2599] avg mini-batch loss: 0.053
[epoch: 1, i: 2699] avg mini-batch loss: 0.042
[epoch: 1, i: 2799] avg mini-batch loss: 0.019
[epoch: 1, i: 2899] avg mini-batch loss: 0.013
[epoch: 1, i: 2999] avg mini-batch loss: 0.021
[epoch: 1, i: 3099] avg mini-batch loss: 0.034
[epoch: 1, i: 3199] avg mini-batch loss: 0.024
[epoch: 1, i: 3299] avg mini-batch loss: 0.043
[epoch: 1, i: 3399] avg mini-batch loss: 0.019
[epoch: 1, i: 3499] avg mini-batch loss: 0.014
[epoch: 1, i: 3599] avg mini-batch loss: 0.013
[epoch: 1, i: 3699] avg mini-batch loss: 0.016
[epoch: 1, i: 3799] avg mini-batch loss: 0.026
[epoch: 1, i: 3899] avg mini-batch loss: 0.030
[epoch: 1, i: 3999] avg mini-batch loss: 0.023
[epoch: 1, i: 4099] avg mini-batch loss: 0.020
[epoch: 1, i: 4199] avg mini-batch loss: 0.014
[epoch: 1, i: 4299] avg mini-batch loss: 0.044
[epoch: 1, i: 4399] avg mini-batch loss: 0.031
[epoch: 1, i: 4499] avg mini-batch loss: 0.045
[epoch: 1, i: 4599] avg mini-batch loss: 0.037
[epoch: 1, i: 4699] avg mini-batch loss: 0.019
[epoch: 1, i: 4799] avg mini-batch loss: 0.019
[epoch: 1, i: 4899] avg mini-batch loss: 0.023
[epoch: 1, i: 4999] avg mini-batch loss: 0.028
[epoch: 1, i: 5099] avg mini-batch loss: 0.019
[epoch: 1, i: 5199] avg mini-batch loss: 0.017
[epoch: 1, i: 5299] avg mini-batch loss: 0.011
[epoch: 1, i: 5399] avg mini-batch loss: 0.051
[epoch: 1, i: 5499] avg mini-batch loss: 0.046
[epoch: 1, i: 5599] avg mini-batch loss: 0.033
[epoch: 1, i: 5699] avg mini-batch loss: 0.051

[epoch: 1, i: 5799] avg mini-batch loss: 0.055
[epoch: 1, i: 5899] avg mini-batch loss: 0.031
[epoch: 1, i: 5999] avg mini-batch loss: 0.039
[epoch: 1, i: 6099] avg mini-batch loss: 0.059
[epoch: 1, i: 6199] avg mini-batch loss: 0.037
[epoch: 1, i: 6299] avg mini-batch loss: 0.060
[epoch: 1, i: 6399] avg mini-batch loss: 0.032
[epoch: 1, i: 6499] avg mini-batch loss: 0.033
[epoch: 1, i: 6599] avg mini-batch loss: 0.060
[epoch: 1, i: 6699] avg mini-batch loss: 0.046
[epoch: 1, i: 6799] avg mini-batch loss: 0.026
[epoch: 1, i: 6899] avg mini-batch loss: 0.050
[epoch: 1, i: 6999] avg mini-batch loss: 0.021
[epoch: 1, i: 7099] avg mini-batch loss: 0.029
[epoch: 1, i: 7199] avg mini-batch loss: 0.041
[epoch: 1, i: 7299] avg mini-batch loss: 0.045
[epoch: 1, i: 7399] avg mini-batch loss: 0.043
[epoch: 1, i: 7499] avg mini-batch loss: 0.026
[epoch: 1, i: 7599] avg mini-batch loss: 0.050
[epoch: 1, i: 7699] avg mini-batch loss: 0.064
[epoch: 1, i: 7799] avg mini-batch loss: 0.030
[epoch: 1, i: 7899] avg mini-batch loss: 0.038
[epoch: 1, i: 7999] avg mini-batch loss: 0.037
[epoch: 1, i: 8099] avg mini-batch loss: 0.014
[epoch: 1, i: 8199] avg mini-batch loss: 0.052
[epoch: 1, i: 8299] avg mini-batch loss: 0.059
[epoch: 1, i: 8399] avg mini-batch loss: 0.045
[epoch: 1, i: 8499] avg mini-batch loss: 0.049
[epoch: 1, i: 8599] avg mini-batch loss: 0.059
[epoch: 1, i: 8699] avg mini-batch loss: 0.088
[epoch: 1, i: 8799] avg mini-batch loss: 0.090
[epoch: 1, i: 8899] avg mini-batch loss: 0.082
[epoch: 1, i: 8999] avg mini-batch loss: 0.088
[epoch: 1, i: 9099] avg mini-batch loss: 0.038
[epoch: 1, i: 9199] avg mini-batch loss: 0.040
[epoch: 1, i: 9299] avg mini-batch loss: 0.025
[epoch: 1, i: 9399] avg mini-batch loss: 0.033
[epoch: 1, i: 9499] avg mini-batch loss: 0.049
[epoch: 1, i: 9599] avg mini-batch loss: 0.080
[epoch: 1, i: 9699] avg mini-batch loss: 0.096
[epoch: 1, i: 9799] avg mini-batch loss: 0.048
[epoch: 1, i: 9899] avg mini-batch loss: 0.103
[epoch: 1, i: 9999] avg mini-batch loss: 0.055
[epoch: 1, i: 10099] avg mini-batch loss: 0.060
[epoch: 1, i: 10199] avg mini-batch loss: 0.072
[epoch: 1, i: 10299] avg mini-batch loss: 0.095
[epoch: 1, i: 10399] avg mini-batch loss: 0.070
[epoch: 1, i: 10499] avg mini-batch loss: 0.026

[epoch: 1, i: 10599] avg mini-batch loss: 0.115
[epoch: 1, i: 10699] avg mini-batch loss: 0.035
[epoch: 1, i: 10799] avg mini-batch loss: 0.056
[epoch: 1, i: 10899] avg mini-batch loss: 0.041
[epoch: 1, i: 10999] avg mini-batch loss: 0.047
[epoch: 1, i: 11099] avg mini-batch loss: 0.027
[epoch: 1, i: 11199] avg mini-batch loss: 0.030
[epoch: 1, i: 11299] avg mini-batch loss: 0.054
[epoch: 1, i: 11399] avg mini-batch loss: 0.015
[epoch: 1, i: 11499] avg mini-batch loss: 0.064
[epoch: 1, i: 11599] avg mini-batch loss: 0.057
[epoch: 1, i: 11699] avg mini-batch loss: 0.066
[epoch: 1, i: 11799] avg mini-batch loss: 0.050
[epoch: 1, i: 11899] avg mini-batch loss: 0.049
[epoch: 1, i: 11999] avg mini-batch loss: 0.069
[epoch: 1, i: 12099] avg mini-batch loss: 0.037
[epoch: 1, i: 12199] avg mini-batch loss: 0.036
[epoch: 1, i: 12299] avg mini-batch loss: 0.072
[epoch: 1, i: 12399] avg mini-batch loss: 0.058
[epoch: 1, i: 12499] avg mini-batch loss: 0.059
[epoch: 2, i: 99] avg mini-batch loss: 0.036
[epoch: 2, i: 199] avg mini-batch loss: 0.046
[epoch: 2, i: 299] avg mini-batch loss: 0.035
[epoch: 2, i: 399] avg mini-batch loss: 0.036
[epoch: 2, i: 499] avg mini-batch loss: 0.021
[epoch: 2, i: 599] avg mini-batch loss: 0.046
[epoch: 2, i: 699] avg mini-batch loss: 0.035
[epoch: 2, i: 799] avg mini-batch loss: 0.011
[epoch: 2, i: 899] avg mini-batch loss: 0.015
[epoch: 2, i: 999] avg mini-batch loss: 0.027
[epoch: 2, i: 1099] avg mini-batch loss: 0.037
[epoch: 2, i: 1199] avg mini-batch loss: 0.089
[epoch: 2, i: 1299] avg mini-batch loss: 0.040
[epoch: 2, i: 1399] avg mini-batch loss: 0.042
[epoch: 2, i: 1499] avg mini-batch loss: 0.071
[epoch: 2, i: 1599] avg mini-batch loss: 0.083
[epoch: 2, i: 1699] avg mini-batch loss: 0.013
[epoch: 2, i: 1799] avg mini-batch loss: 0.028
[epoch: 2, i: 1899] avg mini-batch loss: 0.053
[epoch: 2, i: 1999] avg mini-batch loss: 0.037
[epoch: 2, i: 2099] avg mini-batch loss: 0.049
[epoch: 2, i: 2199] avg mini-batch loss: 0.041
[epoch: 2, i: 2299] avg mini-batch loss: 0.037
[epoch: 2, i: 2399] avg mini-batch loss: 0.019
[epoch: 2, i: 2499] avg mini-batch loss: 0.010
[epoch: 2, i: 2599] avg mini-batch loss: 0.036
[epoch: 2, i: 2699] avg mini-batch loss: 0.029
[epoch: 2, i: 2799] avg mini-batch loss: 0.016

[epoch: 2, i: 2899] avg mini-batch loss: 0.015
[epoch: 2, i: 2999] avg mini-batch loss: 0.037
[epoch: 2, i: 3099] avg mini-batch loss: 0.023
[epoch: 2, i: 3199] avg mini-batch loss: 0.075
[epoch: 2, i: 3299] avg mini-batch loss: 0.032
[epoch: 2, i: 3399] avg mini-batch loss: 0.026
[epoch: 2, i: 3499] avg mini-batch loss: 0.090
[epoch: 2, i: 3599] avg mini-batch loss: 0.035
[epoch: 2, i: 3699] avg mini-batch loss: 0.060
[epoch: 2, i: 3799] avg mini-batch loss: 0.037
[epoch: 2, i: 3899] avg mini-batch loss: 0.045
[epoch: 2, i: 3999] avg mini-batch loss: 0.016
[epoch: 2, i: 4099] avg mini-batch loss: 0.065
[epoch: 2, i: 4199] avg mini-batch loss: 0.051
[epoch: 2, i: 4299] avg mini-batch loss: 0.049
[epoch: 2, i: 4399] avg mini-batch loss: 0.082
[epoch: 2, i: 4499] avg mini-batch loss: 0.060
[epoch: 2, i: 4599] avg mini-batch loss: 0.035
[epoch: 2, i: 4699] avg mini-batch loss: 0.030
[epoch: 2, i: 4799] avg mini-batch loss: 0.037
[epoch: 2, i: 4899] avg mini-batch loss: 0.055
[epoch: 2, i: 4999] avg mini-batch loss: 0.054
[epoch: 2, i: 5099] avg mini-batch loss: 0.077
[epoch: 2, i: 5199] avg mini-batch loss: 0.067
[epoch: 2, i: 5299] avg mini-batch loss: 0.034
[epoch: 2, i: 5399] avg mini-batch loss: 0.024
[epoch: 2, i: 5499] avg mini-batch loss: 0.032
[epoch: 2, i: 5599] avg mini-batch loss: 0.026
[epoch: 2, i: 5699] avg mini-batch loss: 0.023
[epoch: 2, i: 5799] avg mini-batch loss: 0.022
[epoch: 2, i: 5899] avg mini-batch loss: 0.054
[epoch: 2, i: 5999] avg mini-batch loss: 0.031
[epoch: 2, i: 6099] avg mini-batch loss: 0.039
[epoch: 2, i: 6199] avg mini-batch loss: 0.026
[epoch: 2, i: 6299] avg mini-batch loss: 0.025
[epoch: 2, i: 6399] avg mini-batch loss: 0.019
[epoch: 2, i: 6499] avg mini-batch loss: 0.050
[epoch: 2, i: 6599] avg mini-batch loss: 0.041
[epoch: 2, i: 6699] avg mini-batch loss: 0.015
[epoch: 2, i: 6799] avg mini-batch loss: 0.009
[epoch: 2, i: 6899] avg mini-batch loss: 0.066
[epoch: 2, i: 6999] avg mini-batch loss: 0.043
[epoch: 2, i: 7099] avg mini-batch loss: 0.032
[epoch: 2, i: 7199] avg mini-batch loss: 0.031
[epoch: 2, i: 7299] avg mini-batch loss: 0.042
[epoch: 2, i: 7399] avg mini-batch loss: 0.038
[epoch: 2, i: 7499] avg mini-batch loss: 0.033
[epoch: 2, i: 7599] avg mini-batch loss: 0.025

[epoch: 2, i: 7699] avg mini-batch loss: 0.020
[epoch: 2, i: 7799] avg mini-batch loss: 0.010
[epoch: 2, i: 7899] avg mini-batch loss: 0.031
[epoch: 2, i: 7999] avg mini-batch loss: 0.019
[epoch: 2, i: 8099] avg mini-batch loss: 0.021
[epoch: 2, i: 8199] avg mini-batch loss: 0.028
[epoch: 2, i: 8299] avg mini-batch loss: 0.011
[epoch: 2, i: 8399] avg mini-batch loss: 0.016
[epoch: 2, i: 8499] avg mini-batch loss: 0.030
[epoch: 2, i: 8599] avg mini-batch loss: 0.036
[epoch: 2, i: 8699] avg mini-batch loss: 0.016
[epoch: 2, i: 8799] avg mini-batch loss: 0.077
[epoch: 2, i: 8899] avg mini-batch loss: 0.096
[epoch: 2, i: 8999] avg mini-batch loss: 0.048
[epoch: 2, i: 9099] avg mini-batch loss: 0.051
[epoch: 2, i: 9199] avg mini-batch loss: 0.056
[epoch: 2, i: 9299] avg mini-batch loss: 0.069
[epoch: 2, i: 9399] avg mini-batch loss: 0.050
[epoch: 2, i: 9499] avg mini-batch loss: 0.034
[epoch: 2, i: 9599] avg mini-batch loss: 0.044
[epoch: 2, i: 9699] avg mini-batch loss: 0.019
[epoch: 2, i: 9799] avg mini-batch loss: 0.045
[epoch: 2, i: 9899] avg mini-batch loss: 0.032
[epoch: 2, i: 9999] avg mini-batch loss: 0.016
[epoch: 2, i: 10099] avg mini-batch loss: 0.014
[epoch: 2, i: 10199] avg mini-batch loss: 0.021
[epoch: 2, i: 10299] avg mini-batch loss: 0.061
[epoch: 2, i: 10399] avg mini-batch loss: 0.047
[epoch: 2, i: 10499] avg mini-batch loss: 0.022
[epoch: 2, i: 10599] avg mini-batch loss: 0.031
[epoch: 2, i: 10699] avg mini-batch loss: 0.031
[epoch: 2, i: 10799] avg mini-batch loss: 0.013
[epoch: 2, i: 10899] avg mini-batch loss: 0.039
[epoch: 2, i: 10999] avg mini-batch loss: 0.033
[epoch: 2, i: 11099] avg mini-batch loss: 0.059
[epoch: 2, i: 11199] avg mini-batch loss: 0.054
[epoch: 2, i: 11299] avg mini-batch loss: 0.080
[epoch: 2, i: 11399] avg mini-batch loss: 0.038
[epoch: 2, i: 11499] avg mini-batch loss: 0.044
[epoch: 2, i: 11599] avg mini-batch loss: 0.086
[epoch: 2, i: 11699] avg mini-batch loss: 0.054
[epoch: 2, i: 11799] avg mini-batch loss: 0.074
[epoch: 2, i: 11899] avg mini-batch loss: 0.059
[epoch: 2, i: 11999] avg mini-batch loss: 0.027
[epoch: 2, i: 12099] avg mini-batch loss: 0.052
[epoch: 2, i: 12199] avg mini-batch loss: 0.029
[epoch: 2, i: 12299] avg mini-batch loss: 0.059
[epoch: 2, i: 12399] avg mini-batch loss: 0.044

[epoch: 2, i: 12499] avg mini-batch loss: 0.071
[epoch: 3, i: 99] avg mini-batch loss: 0.039
[epoch: 3, i: 199] avg mini-batch loss: 0.033
[epoch: 3, i: 299] avg mini-batch loss: 0.008
[epoch: 3, i: 399] avg mini-batch loss: 0.015
[epoch: 3, i: 499] avg mini-batch loss: 0.027
[epoch: 3, i: 599] avg mini-batch loss: 0.023
[epoch: 3, i: 699] avg mini-batch loss: 0.016
[epoch: 3, i: 799] avg mini-batch loss: 0.017
[epoch: 3, i: 899] avg mini-batch loss: 0.064
[epoch: 3, i: 999] avg mini-batch loss: 0.048
[epoch: 3, i: 1099] avg mini-batch loss: 0.068
[epoch: 3, i: 1199] avg mini-batch loss: 0.059
[epoch: 3, i: 1299] avg mini-batch loss: 0.014
[epoch: 3, i: 1399] avg mini-batch loss: 0.014
[epoch: 3, i: 1499] avg mini-batch loss: 0.084
[epoch: 3, i: 1599] avg mini-batch loss: 0.028
[epoch: 3, i: 1699] avg mini-batch loss: 0.055
[epoch: 3, i: 1799] avg mini-batch loss: 0.034
[epoch: 3, i: 1899] avg mini-batch loss: 0.085
[epoch: 3, i: 1999] avg mini-batch loss: 0.046
[epoch: 3, i: 2099] avg mini-batch loss: 0.075
[epoch: 3, i: 2199] avg mini-batch loss: 0.069
[epoch: 3, i: 2299] avg mini-batch loss: 0.035
[epoch: 3, i: 2399] avg mini-batch loss: 0.036
[epoch: 3, i: 2499] avg mini-batch loss: 0.055
[epoch: 3, i: 2599] avg mini-batch loss: 0.033
[epoch: 3, i: 2699] avg mini-batch loss: 0.066
[epoch: 3, i: 2799] avg mini-batch loss: 0.055
[epoch: 3, i: 2899] avg mini-batch loss: 0.039
[epoch: 3, i: 2999] avg mini-batch loss: 0.027
[epoch: 3, i: 3099] avg mini-batch loss: 0.043
[epoch: 3, i: 3199] avg mini-batch loss: 0.012
[epoch: 3, i: 3299] avg mini-batch loss: 0.009
[epoch: 3, i: 3399] avg mini-batch loss: 0.045
[epoch: 3, i: 3499] avg mini-batch loss: 0.024
[epoch: 3, i: 3599] avg mini-batch loss: 0.037
[epoch: 3, i: 3699] avg mini-batch loss: 0.071
[epoch: 3, i: 3799] avg mini-batch loss: 0.030
[epoch: 3, i: 3899] avg mini-batch loss: 0.047
[epoch: 3, i: 3999] avg mini-batch loss: 0.050
[epoch: 3, i: 4099] avg mini-batch loss: 0.037
[epoch: 3, i: 4199] avg mini-batch loss: 0.068
[epoch: 3, i: 4299] avg mini-batch loss: 0.024
[epoch: 3, i: 4399] avg mini-batch loss: 0.057
[epoch: 3, i: 4499] avg mini-batch loss: 0.030
[epoch: 3, i: 4599] avg mini-batch loss: 0.049
[epoch: 3, i: 4699] avg mini-batch loss: 0.034

[epoch: 3, i: 4799] avg mini-batch loss: 0.034
[epoch: 3, i: 4899] avg mini-batch loss: 0.018
[epoch: 3, i: 4999] avg mini-batch loss: 0.028
[epoch: 3, i: 5099] avg mini-batch loss: 0.039
[epoch: 3, i: 5199] avg mini-batch loss: 0.039
[epoch: 3, i: 5299] avg mini-batch loss: 0.051
[epoch: 3, i: 5399] avg mini-batch loss: 0.067
[epoch: 3, i: 5499] avg mini-batch loss: 0.029
[epoch: 3, i: 5599] avg mini-batch loss: 0.024
[epoch: 3, i: 5699] avg mini-batch loss: 0.068
[epoch: 3, i: 5799] avg mini-batch loss: 0.070
[epoch: 3, i: 5899] avg mini-batch loss: 0.076
[epoch: 3, i: 5999] avg mini-batch loss: 0.067
[epoch: 3, i: 6099] avg mini-batch loss: 0.019
[epoch: 3, i: 6199] avg mini-batch loss: 0.054
[epoch: 3, i: 6299] avg mini-batch loss: 0.081
[epoch: 3, i: 6399] avg mini-batch loss: 0.071
[epoch: 3, i: 6499] avg mini-batch loss: 0.036
[epoch: 3, i: 6599] avg mini-batch loss: 0.025
[epoch: 3, i: 6699] avg mini-batch loss: 0.034
[epoch: 3, i: 6799] avg mini-batch loss: 0.020
[epoch: 3, i: 6899] avg mini-batch loss: 0.049
[epoch: 3, i: 6999] avg mini-batch loss: 0.040
[epoch: 3, i: 7099] avg mini-batch loss: 0.021
[epoch: 3, i: 7199] avg mini-batch loss: 0.053
[epoch: 3, i: 7299] avg mini-batch loss: 0.022
[epoch: 3, i: 7399] avg mini-batch loss: 0.033
[epoch: 3, i: 7499] avg mini-batch loss: 0.063
[epoch: 3, i: 7599] avg mini-batch loss: 0.086
[epoch: 3, i: 7699] avg mini-batch loss: 0.068
[epoch: 3, i: 7799] avg mini-batch loss: 0.048
[epoch: 3, i: 7899] avg mini-batch loss: 0.087
[epoch: 3, i: 7999] avg mini-batch loss: 0.074
[epoch: 3, i: 8099] avg mini-batch loss: 0.087
[epoch: 3, i: 8199] avg mini-batch loss: 0.054
[epoch: 3, i: 8299] avg mini-batch loss: 0.101
[epoch: 3, i: 8399] avg mini-batch loss: 0.065
[epoch: 3, i: 8499] avg mini-batch loss: 0.066
[epoch: 3, i: 8599] avg mini-batch loss: 0.032
[epoch: 3, i: 8699] avg mini-batch loss: 0.095
[epoch: 3, i: 8799] avg mini-batch loss: 0.098
[epoch: 3, i: 8899] avg mini-batch loss: 0.076
[epoch: 3, i: 8999] avg mini-batch loss: 0.039
[epoch: 3, i: 9099] avg mini-batch loss: 0.032
[epoch: 3, i: 9199] avg mini-batch loss: 0.049
[epoch: 3, i: 9299] avg mini-batch loss: 0.033
[epoch: 3, i: 9399] avg mini-batch loss: 0.034
[epoch: 3, i: 9499] avg mini-batch loss: 0.011

[epoch: 3, i: 9599] avg mini-batch loss: 0.015
[epoch: 3, i: 9699] avg mini-batch loss: 0.028
[epoch: 3, i: 9799] avg mini-batch loss: 0.032
[epoch: 3, i: 9899] avg mini-batch loss: 0.044
[epoch: 3, i: 9999] avg mini-batch loss: 0.058
[epoch: 3, i: 10099] avg mini-batch loss: 0.061
[epoch: 3, i: 10199] avg mini-batch loss: 0.054
[epoch: 3, i: 10299] avg mini-batch loss: 0.127
[epoch: 3, i: 10399] avg mini-batch loss: 0.073
[epoch: 3, i: 10499] avg mini-batch loss: 0.051
[epoch: 3, i: 10599] avg mini-batch loss: 0.026
[epoch: 3, i: 10699] avg mini-batch loss: 0.056
[epoch: 3, i: 10799] avg mini-batch loss: 0.054
[epoch: 3, i: 10899] avg mini-batch loss: 0.079
[epoch: 3, i: 10999] avg mini-batch loss: 0.090
[epoch: 3, i: 11099] avg mini-batch loss: 0.031
[epoch: 3, i: 11199] avg mini-batch loss: 0.029
[epoch: 3, i: 11299] avg mini-batch loss: 0.062
[epoch: 3, i: 11399] avg mini-batch loss: 0.033
[epoch: 3, i: 11499] avg mini-batch loss: 0.052
[epoch: 3, i: 11599] avg mini-batch loss: 0.042
[epoch: 3, i: 11699] avg mini-batch loss: 0.024
[epoch: 3, i: 11799] avg mini-batch loss: 0.042
[epoch: 3, i: 11899] avg mini-batch loss: 0.061
[epoch: 3, i: 11999] avg mini-batch loss: 0.060
[epoch: 3, i: 12099] avg mini-batch loss: 0.075
[epoch: 3, i: 12199] avg mini-batch loss: 0.056
[epoch: 3, i: 12299] avg mini-batch loss: 0.052
[epoch: 3, i: 12399] avg mini-batch loss: 0.029
[epoch: 3, i: 12499] avg mini-batch loss: 0.037
[epoch: 4, i: 99] avg mini-batch loss: 0.028
[epoch: 4, i: 199] avg mini-batch loss: 0.020
[epoch: 4, i: 299] avg mini-batch loss: 0.013
[epoch: 4, i: 399] avg mini-batch loss: 0.019
[epoch: 4, i: 499] avg mini-batch loss: 0.012
[epoch: 4, i: 599] avg mini-batch loss: 0.021
[epoch: 4, i: 699] avg mini-batch loss: 0.019
[epoch: 4, i: 799] avg mini-batch loss: 0.010
[epoch: 4, i: 899] avg mini-batch loss: 0.014
[epoch: 4, i: 999] avg mini-batch loss: 0.008
[epoch: 4, i: 1099] avg mini-batch loss: 0.007
[epoch: 4, i: 1199] avg mini-batch loss: 0.010
[epoch: 4, i: 1299] avg mini-batch loss: 0.016
[epoch: 4, i: 1399] avg mini-batch loss: 0.014
[epoch: 4, i: 1499] avg mini-batch loss: 0.034
[epoch: 4, i: 1599] avg mini-batch loss: 0.014
[epoch: 4, i: 1699] avg mini-batch loss: 0.015
[epoch: 4, i: 1799] avg mini-batch loss: 0.005

[epoch: 4, i: 1899] avg mini-batch loss: 0.023
[epoch: 4, i: 1999] avg mini-batch loss: 0.014
[epoch: 4, i: 2099] avg mini-batch loss: 0.030
[epoch: 4, i: 2199] avg mini-batch loss: 0.023
[epoch: 4, i: 2299] avg mini-batch loss: 0.033
[epoch: 4, i: 2399] avg mini-batch loss: 0.019
[epoch: 4, i: 2499] avg mini-batch loss: 0.034
[epoch: 4, i: 2599] avg mini-batch loss: 0.028
[epoch: 4, i: 2699] avg mini-batch loss: 0.046
[epoch: 4, i: 2799] avg mini-batch loss: 0.019
[epoch: 4, i: 2899] avg mini-batch loss: 0.045
[epoch: 4, i: 2999] avg mini-batch loss: 0.031
[epoch: 4, i: 3099] avg mini-batch loss: 0.043
[epoch: 4, i: 3199] avg mini-batch loss: 0.035
[epoch: 4, i: 3299] avg mini-batch loss: 0.039
[epoch: 4, i: 3399] avg mini-batch loss: 0.034
[epoch: 4, i: 3499] avg mini-batch loss: 0.014
[epoch: 4, i: 3599] avg mini-batch loss: 0.016
[epoch: 4, i: 3699] avg mini-batch loss: 0.011
[epoch: 4, i: 3799] avg mini-batch loss: 0.009
[epoch: 4, i: 3899] avg mini-batch loss: 0.017
[epoch: 4, i: 3999] avg mini-batch loss: 0.024
[epoch: 4, i: 4099] avg mini-batch loss: 0.029
[epoch: 4, i: 4199] avg mini-batch loss: 0.021
[epoch: 4, i: 4299] avg mini-batch loss: 0.016
[epoch: 4, i: 4399] avg mini-batch loss: 0.014
[epoch: 4, i: 4499] avg mini-batch loss: 0.049
[epoch: 4, i: 4599] avg mini-batch loss: 0.027
[epoch: 4, i: 4699] avg mini-batch loss: 0.030
[epoch: 4, i: 4799] avg mini-batch loss: 0.025
[epoch: 4, i: 4899] avg mini-batch loss: 0.039
[epoch: 4, i: 4999] avg mini-batch loss: 0.026
[epoch: 4, i: 5099] avg mini-batch loss: 0.028
[epoch: 4, i: 5199] avg mini-batch loss: 0.035
[epoch: 4, i: 5299] avg mini-batch loss: 0.050
[epoch: 4, i: 5399] avg mini-batch loss: 0.022
[epoch: 4, i: 5499] avg mini-batch loss: 0.043
[epoch: 4, i: 5599] avg mini-batch loss: 0.072
[epoch: 4, i: 5699] avg mini-batch loss: 0.041
[epoch: 4, i: 5799] avg mini-batch loss: 0.040
[epoch: 4, i: 5899] avg mini-batch loss: 0.017
[epoch: 4, i: 5999] avg mini-batch loss: 0.021
[epoch: 4, i: 6099] avg mini-batch loss: 0.067
[epoch: 4, i: 6199] avg mini-batch loss: 0.033
[epoch: 4, i: 6299] avg mini-batch loss: 0.023
[epoch: 4, i: 6399] avg mini-batch loss: 0.026
[epoch: 4, i: 6499] avg mini-batch loss: 0.020
[epoch: 4, i: 6599] avg mini-batch loss: 0.052

[epoch: 4, i: 6699] avg mini-batch loss: 0.063
[epoch: 4, i: 6799] avg mini-batch loss: 0.019
[epoch: 4, i: 6899] avg mini-batch loss: 0.015
[epoch: 4, i: 6999] avg mini-batch loss: 0.009
[epoch: 4, i: 7099] avg mini-batch loss: 0.029
[epoch: 4, i: 7199] avg mini-batch loss: 0.062
[epoch: 4, i: 7299] avg mini-batch loss: 0.105
[epoch: 4, i: 7399] avg mini-batch loss: 0.032
[epoch: 4, i: 7499] avg mini-batch loss: 0.057
[epoch: 4, i: 7599] avg mini-batch loss: 0.052
[epoch: 4, i: 7699] avg mini-batch loss: 0.025
[epoch: 4, i: 7799] avg mini-batch loss: 0.071
[epoch: 4, i: 7899] avg mini-batch loss: 0.032
[epoch: 4, i: 7999] avg mini-batch loss: 0.030
[epoch: 4, i: 8099] avg mini-batch loss: 0.019
[epoch: 4, i: 8199] avg mini-batch loss: 0.037
[epoch: 4, i: 8299] avg mini-batch loss: 0.044
[epoch: 4, i: 8399] avg mini-batch loss: 0.045
[epoch: 4, i: 8499] avg mini-batch loss: 0.026
[epoch: 4, i: 8599] avg mini-batch loss: 0.036
[epoch: 4, i: 8699] avg mini-batch loss: 0.057
[epoch: 4, i: 8799] avg mini-batch loss: 0.037
[epoch: 4, i: 8899] avg mini-batch loss: 0.038
[epoch: 4, i: 8999] avg mini-batch loss: 0.030
[epoch: 4, i: 9099] avg mini-batch loss: 0.018
[epoch: 4, i: 9199] avg mini-batch loss: 0.024
[epoch: 4, i: 9299] avg mini-batch loss: 0.025
[epoch: 4, i: 9399] avg mini-batch loss: 0.032
[epoch: 4, i: 9499] avg mini-batch loss: 0.024
[epoch: 4, i: 9599] avg mini-batch loss: 0.048
[epoch: 4, i: 9699] avg mini-batch loss: 0.019
[epoch: 4, i: 9799] avg mini-batch loss: 0.011
[epoch: 4, i: 9899] avg mini-batch loss: 0.028
[epoch: 4, i: 9999] avg mini-batch loss: 0.027
[epoch: 4, i: 10099] avg mini-batch loss: 0.013
[epoch: 4, i: 10199] avg mini-batch loss: 0.018
[epoch: 4, i: 10299] avg mini-batch loss: 0.028
[epoch: 4, i: 10399] avg mini-batch loss: 0.027
[epoch: 4, i: 10499] avg mini-batch loss: 0.057
[epoch: 4, i: 10599] avg mini-batch loss: 0.034
[epoch: 4, i: 10699] avg mini-batch loss: 0.071
[epoch: 4, i: 10799] avg mini-batch loss: 0.033
[epoch: 4, i: 10899] avg mini-batch loss: 0.023
[epoch: 4, i: 10999] avg mini-batch loss: 0.022
[epoch: 4, i: 11099] avg mini-batch loss: 0.032
[epoch: 4, i: 11199] avg mini-batch loss: 0.054
[epoch: 4, i: 11299] avg mini-batch loss: 0.078
[epoch: 4, i: 11399] avg mini-batch loss: 0.099

[epoch: 4, i: 11499] avg mini-batch loss: 0.038
[epoch: 4, i: 11599] avg mini-batch loss: 0.077
[epoch: 4, i: 11699] avg mini-batch loss: 0.087
[epoch: 4, i: 11799] avg mini-batch loss: 0.042
[epoch: 4, i: 11899] avg mini-batch loss: 0.087
[epoch: 4, i: 11999] avg mini-batch loss: 0.068
[epoch: 4, i: 12099] avg mini-batch loss: 0.084
[epoch: 4, i: 12199] avg mini-batch loss: 0.051
[epoch: 4, i: 12299] avg mini-batch loss: 0.045
[epoch: 4, i: 12399] avg mini-batch loss: 0.033
[epoch: 4, i: 12499] avg mini-batch loss: 0.047
[epoch: 5, i: 99] avg mini-batch loss: 0.032
[epoch: 5, i: 199] avg mini-batch loss: 0.039
[epoch: 5, i: 299] avg mini-batch loss: 0.021
[epoch: 5, i: 399] avg mini-batch loss: 0.012
[epoch: 5, i: 499] avg mini-batch loss: 0.015
[epoch: 5, i: 599] avg mini-batch loss: 0.005
[epoch: 5, i: 699] avg mini-batch loss: 0.015
[epoch: 5, i: 799] avg mini-batch loss: 0.010
[epoch: 5, i: 899] avg mini-batch loss: 0.010
[epoch: 5, i: 999] avg mini-batch loss: 0.006
[epoch: 5, i: 1099] avg mini-batch loss: 0.027
[epoch: 5, i: 1199] avg mini-batch loss: 0.017
[epoch: 5, i: 1299] avg mini-batch loss: 0.014
[epoch: 5, i: 1399] avg mini-batch loss: 0.036
[epoch: 5, i: 1499] avg mini-batch loss: 0.031
[epoch: 5, i: 1599] avg mini-batch loss: 0.018
[epoch: 5, i: 1699] avg mini-batch loss: 0.051
[epoch: 5, i: 1799] avg mini-batch loss: 0.023
[epoch: 5, i: 1899] avg mini-batch loss: 0.008
[epoch: 5, i: 1999] avg mini-batch loss: 0.015
[epoch: 5, i: 2099] avg mini-batch loss: 0.004
[epoch: 5, i: 2199] avg mini-batch loss: 0.011
[epoch: 5, i: 2299] avg mini-batch loss: 0.024
[epoch: 5, i: 2399] avg mini-batch loss: 0.035
[epoch: 5, i: 2499] avg mini-batch loss: 0.029
[epoch: 5, i: 2599] avg mini-batch loss: 0.024
[epoch: 5, i: 2699] avg mini-batch loss: 0.018
[epoch: 5, i: 2799] avg mini-batch loss: 0.012
[epoch: 5, i: 2899] avg mini-batch loss: 0.010
[epoch: 5, i: 2999] avg mini-batch loss: 0.020
[epoch: 5, i: 3099] avg mini-batch loss: 0.066
[epoch: 5, i: 3199] avg mini-batch loss: 0.038
[epoch: 5, i: 3299] avg mini-batch loss: 0.014
[epoch: 5, i: 3399] avg mini-batch loss: 0.043
[epoch: 5, i: 3499] avg mini-batch loss: 0.020
[epoch: 5, i: 3599] avg mini-batch loss: 0.039
[epoch: 5, i: 3699] avg mini-batch loss: 0.052

[epoch: 5, i: 3799] avg mini-batch loss: 0.086
[epoch: 5, i: 3899] avg mini-batch loss: 0.041
[epoch: 5, i: 3999] avg mini-batch loss: 0.029
[epoch: 5, i: 4099] avg mini-batch loss: 0.027
[epoch: 5, i: 4199] avg mini-batch loss: 0.024
[epoch: 5, i: 4299] avg mini-batch loss: 0.008
[epoch: 5, i: 4399] avg mini-batch loss: 0.015
[epoch: 5, i: 4499] avg mini-batch loss: 0.016
[epoch: 5, i: 4599] avg mini-batch loss: 0.030
[epoch: 5, i: 4699] avg mini-batch loss: 0.047
[epoch: 5, i: 4799] avg mini-batch loss: 0.026
[epoch: 5, i: 4899] avg mini-batch loss: 0.041
[epoch: 5, i: 4999] avg mini-batch loss: 0.036
[epoch: 5, i: 5099] avg mini-batch loss: 0.019
[epoch: 5, i: 5199] avg mini-batch loss: 0.044
[epoch: 5, i: 5299] avg mini-batch loss: 0.023
[epoch: 5, i: 5399] avg mini-batch loss: 0.038
[epoch: 5, i: 5499] avg mini-batch loss: 0.042
[epoch: 5, i: 5599] avg mini-batch loss: 0.043
[epoch: 5, i: 5699] avg mini-batch loss: 0.025
[epoch: 5, i: 5799] avg mini-batch loss: 0.026
[epoch: 5, i: 5899] avg mini-batch loss: 0.014
[epoch: 5, i: 5999] avg mini-batch loss: 0.013
[epoch: 5, i: 6099] avg mini-batch loss: 0.059
[epoch: 5, i: 6199] avg mini-batch loss: 0.022
[epoch: 5, i: 6299] avg mini-batch loss: 0.008
[epoch: 5, i: 6399] avg mini-batch loss: 0.039
[epoch: 5, i: 6499] avg mini-batch loss: 0.018
[epoch: 5, i: 6599] avg mini-batch loss: 0.024
[epoch: 5, i: 6699] avg mini-batch loss: 0.044
[epoch: 5, i: 6799] avg mini-batch loss: 0.032
[epoch: 5, i: 6899] avg mini-batch loss: 0.033
[epoch: 5, i: 6999] avg mini-batch loss: 0.024
[epoch: 5, i: 7099] avg mini-batch loss: 0.017
[epoch: 5, i: 7199] avg mini-batch loss: 0.012
[epoch: 5, i: 7299] avg mini-batch loss: 0.033
[epoch: 5, i: 7399] avg mini-batch loss: 0.029
[epoch: 5, i: 7499] avg mini-batch loss: 0.036
[epoch: 5, i: 7599] avg mini-batch loss: 0.008
[epoch: 5, i: 7699] avg mini-batch loss: 0.021
[epoch: 5, i: 7799] avg mini-batch loss: 0.024
[epoch: 5, i: 7899] avg mini-batch loss: 0.018
[epoch: 5, i: 7999] avg mini-batch loss: 0.039
[epoch: 5, i: 8099] avg mini-batch loss: 0.010
[epoch: 5, i: 8199] avg mini-batch loss: 0.016
[epoch: 5, i: 8299] avg mini-batch loss: 0.008
[epoch: 5, i: 8399] avg mini-batch loss: 0.018
[epoch: 5, i: 8499] avg mini-batch loss: 0.017

[epoch: 5, i: 8599] avg mini-batch loss: 0.032
[epoch: 5, i: 8699] avg mini-batch loss: 0.033
[epoch: 5, i: 8799] avg mini-batch loss: 0.028
[epoch: 5, i: 8899] avg mini-batch loss: 0.023
[epoch: 5, i: 8999] avg mini-batch loss: 0.027
[epoch: 5, i: 9099] avg mini-batch loss: 0.045
[epoch: 5, i: 9199] avg mini-batch loss: 0.033
[epoch: 5, i: 9299] avg mini-batch loss: 0.026
[epoch: 5, i: 9399] avg mini-batch loss: 0.019
[epoch: 5, i: 9499] avg mini-batch loss: 0.013
[epoch: 5, i: 9599] avg mini-batch loss: 0.012
[epoch: 5, i: 9699] avg mini-batch loss: 0.043
[epoch: 5, i: 9799] avg mini-batch loss: 0.036
[epoch: 5, i: 9899] avg mini-batch loss: 0.028
[epoch: 5, i: 9999] avg mini-batch loss: 0.013
[epoch: 5, i: 10099] avg mini-batch loss: 0.055
[epoch: 5, i: 10199] avg mini-batch loss: 0.043
[epoch: 5, i: 10299] avg mini-batch loss: 0.050
[epoch: 5, i: 10399] avg mini-batch loss: 0.027
[epoch: 5, i: 10499] avg mini-batch loss: 0.057
[epoch: 5, i: 10599] avg mini-batch loss: 0.033
[epoch: 5, i: 10699] avg mini-batch loss: 0.027
[epoch: 5, i: 10799] avg mini-batch loss: 0.029
[epoch: 5, i: 10899] avg mini-batch loss: 0.021
[epoch: 5, i: 10999] avg mini-batch loss: 0.024
[epoch: 5, i: 11099] avg mini-batch loss: 0.022
[epoch: 5, i: 11199] avg mini-batch loss: 0.078
[epoch: 5, i: 11299] avg mini-batch loss: 0.024
[epoch: 5, i: 11399] avg mini-batch loss: 0.053
[epoch: 5, i: 11499] avg mini-batch loss: 0.040
[epoch: 5, i: 11599] avg mini-batch loss: 0.035
[epoch: 5, i: 11699] avg mini-batch loss: 0.065
[epoch: 5, i: 11799] avg mini-batch loss: 0.038
[epoch: 5, i: 11899] avg mini-batch loss: 0.021
[epoch: 5, i: 11999] avg mini-batch loss: 0.029
[epoch: 5, i: 12099] avg mini-batch loss: 0.016
[epoch: 5, i: 12199] avg mini-batch loss: 0.026
[epoch: 5, i: 12299] avg mini-batch loss: 0.012
[epoch: 5, i: 12399] avg mini-batch loss: 0.028
[epoch: 5, i: 12499] avg mini-batch loss: 0.037
[epoch: 6, i: 99] avg mini-batch loss: 0.023
[epoch: 6, i: 199] avg mini-batch loss: 0.027
[epoch: 6, i: 299] avg mini-batch loss: 0.020
[epoch: 6, i: 399] avg mini-batch loss: 0.036
[epoch: 6, i: 499] avg mini-batch loss: 0.034
[epoch: 6, i: 599] avg mini-batch loss: 0.013
[epoch: 6, i: 699] avg mini-batch loss: 0.017
[epoch: 6, i: 799] avg mini-batch loss: 0.010

[epoch: 6, i: 899] avg mini-batch loss: 0.016
[epoch: 6, i: 999] avg mini-batch loss: 0.004
[epoch: 6, i: 1099] avg mini-batch loss: 0.008
[epoch: 6, i: 1199] avg mini-batch loss: 0.046
[epoch: 6, i: 1299] avg mini-batch loss: 0.023
[epoch: 6, i: 1399] avg mini-batch loss: 0.018
[epoch: 6, i: 1499] avg mini-batch loss: 0.010
[epoch: 6, i: 1599] avg mini-batch loss: 0.017
[epoch: 6, i: 1699] avg mini-batch loss: 0.011
[epoch: 6, i: 1799] avg mini-batch loss: 0.007
[epoch: 6, i: 1899] avg mini-batch loss: 0.013
[epoch: 6, i: 1999] avg mini-batch loss: 0.010
[epoch: 6, i: 2099] avg mini-batch loss: 0.005
[epoch: 6, i: 2199] avg mini-batch loss: 0.016
[epoch: 6, i: 2299] avg mini-batch loss: 0.009
[epoch: 6, i: 2399] avg mini-batch loss: 0.016
[epoch: 6, i: 2499] avg mini-batch loss: 0.012
[epoch: 6, i: 2599] avg mini-batch loss: 0.021
[epoch: 6, i: 2699] avg mini-batch loss: 0.027
[epoch: 6, i: 2799] avg mini-batch loss: 0.029
[epoch: 6, i: 2899] avg mini-batch loss: 0.015
[epoch: 6, i: 2999] avg mini-batch loss: 0.020
[epoch: 6, i: 3099] avg mini-batch loss: 0.005
[epoch: 6, i: 3199] avg mini-batch loss: 0.022
[epoch: 6, i: 3299] avg mini-batch loss: 0.046
[epoch: 6, i: 3399] avg mini-batch loss: 0.026
[epoch: 6, i: 3499] avg mini-batch loss: 0.018
[epoch: 6, i: 3599] avg mini-batch loss: 0.022
[epoch: 6, i: 3699] avg mini-batch loss: 0.017
[epoch: 6, i: 3799] avg mini-batch loss: 0.019
[epoch: 6, i: 3899] avg mini-batch loss: 0.012
[epoch: 6, i: 3999] avg mini-batch loss: 0.032
[epoch: 6, i: 4099] avg mini-batch loss: 0.025
[epoch: 6, i: 4199] avg mini-batch loss: 0.017
[epoch: 6, i: 4299] avg mini-batch loss: 0.015
[epoch: 6, i: 4399] avg mini-batch loss: 0.012
[epoch: 6, i: 4499] avg mini-batch loss: 0.020
[epoch: 6, i: 4599] avg mini-batch loss: 0.025
[epoch: 6, i: 4699] avg mini-batch loss: 0.020
[epoch: 6, i: 4799] avg mini-batch loss: 0.007
[epoch: 6, i: 4899] avg mini-batch loss: 0.022
[epoch: 6, i: 4999] avg mini-batch loss: 0.004
[epoch: 6, i: 5099] avg mini-batch loss: 0.017
[epoch: 6, i: 5199] avg mini-batch loss: 0.011
[epoch: 6, i: 5299] avg mini-batch loss: 0.003
[epoch: 6, i: 5399] avg mini-batch loss: 0.015
[epoch: 6, i: 5499] avg mini-batch loss: 0.002
[epoch: 6, i: 5599] avg mini-batch loss: 0.006

[epoch: 6, i: 5699] avg mini-batch loss: 0.021
[epoch: 6, i: 5799] avg mini-batch loss: 0.026
[epoch: 6, i: 5899] avg mini-batch loss: 0.008
[epoch: 6, i: 5999] avg mini-batch loss: 0.019
[epoch: 6, i: 6099] avg mini-batch loss: 0.013
[epoch: 6, i: 6199] avg mini-batch loss: 0.021
[epoch: 6, i: 6299] avg mini-batch loss: 0.033
[epoch: 6, i: 6399] avg mini-batch loss: 0.024
[epoch: 6, i: 6499] avg mini-batch loss: 0.036
[epoch: 6, i: 6599] avg mini-batch loss: 0.062
[epoch: 6, i: 6699] avg mini-batch loss: 0.075
[epoch: 6, i: 6799] avg mini-batch loss: 0.039
[epoch: 6, i: 6899] avg mini-batch loss: 0.041
[epoch: 6, i: 6999] avg mini-batch loss: 0.019
[epoch: 6, i: 7099] avg mini-batch loss: 0.040
[epoch: 6, i: 7199] avg mini-batch loss: 0.083
[epoch: 6, i: 7299] avg mini-batch loss: 0.058
[epoch: 6, i: 7399] avg mini-batch loss: 0.015
[epoch: 6, i: 7499] avg mini-batch loss: 0.006
[epoch: 6, i: 7599] avg mini-batch loss: 0.027
[epoch: 6, i: 7699] avg mini-batch loss: 0.045
[epoch: 6, i: 7799] avg mini-batch loss: 0.074
[epoch: 6, i: 7899] avg mini-batch loss: 0.026
[epoch: 6, i: 7999] avg mini-batch loss: 0.059
[epoch: 6, i: 8099] avg mini-batch loss: 0.035
[epoch: 6, i: 8199] avg mini-batch loss: 0.040
[epoch: 6, i: 8299] avg mini-batch loss: 0.037
[epoch: 6, i: 8399] avg mini-batch loss: 0.081
[epoch: 6, i: 8499] avg mini-batch loss: 0.044
[epoch: 6, i: 8599] avg mini-batch loss: 0.032
[epoch: 6, i: 8699] avg mini-batch loss: 0.030
[epoch: 6, i: 8799] avg mini-batch loss: 0.028
[epoch: 6, i: 8899] avg mini-batch loss: 0.007
[epoch: 6, i: 8999] avg mini-batch loss: 0.013
[epoch: 6, i: 9099] avg mini-batch loss: 0.010
[epoch: 6, i: 9199] avg mini-batch loss: 0.011
[epoch: 6, i: 9299] avg mini-batch loss: 0.033
[epoch: 6, i: 9399] avg mini-batch loss: 0.018
[epoch: 6, i: 9499] avg mini-batch loss: 0.010
[epoch: 6, i: 9599] avg mini-batch loss: 0.014
[epoch: 6, i: 9699] avg mini-batch loss: 0.038
[epoch: 6, i: 9799] avg mini-batch loss: 0.022
[epoch: 6, i: 9899] avg mini-batch loss: 0.033
[epoch: 6, i: 9999] avg mini-batch loss: 0.046
[epoch: 6, i: 10099] avg mini-batch loss: 0.040
[epoch: 6, i: 10199] avg mini-batch loss: 0.018
[epoch: 6, i: 10299] avg mini-batch loss: 0.062
[epoch: 6, i: 10399] avg mini-batch loss: 0.075

[epoch: 6, i: 10499] avg mini-batch loss: 0.043
[epoch: 6, i: 10599] avg mini-batch loss: 0.043
[epoch: 6, i: 10699] avg mini-batch loss: 0.032
[epoch: 6, i: 10799] avg mini-batch loss: 0.029
[epoch: 6, i: 10899] avg mini-batch loss: 0.014
[epoch: 6, i: 10999] avg mini-batch loss: 0.027
[epoch: 6, i: 11099] avg mini-batch loss: 0.026
[epoch: 6, i: 11199] avg mini-batch loss: 0.013
[epoch: 6, i: 11299] avg mini-batch loss: 0.040
[epoch: 6, i: 11399] avg mini-batch loss: 0.017
[epoch: 6, i: 11499] avg mini-batch loss: 0.010
[epoch: 6, i: 11599] avg mini-batch loss: 0.016
[epoch: 6, i: 11699] avg mini-batch loss: 0.021
[epoch: 6, i: 11799] avg mini-batch loss: 0.022
[epoch: 6, i: 11899] avg mini-batch loss: 0.040
[epoch: 6, i: 11999] avg mini-batch loss: 0.028
[epoch: 6, i: 12099] avg mini-batch loss: 0.048
[epoch: 6, i: 12199] avg mini-batch loss: 0.007
[epoch: 6, i: 12299] avg mini-batch loss: 0.041
[epoch: 6, i: 12399] avg mini-batch loss: 0.036
[epoch: 6, i: 12499] avg mini-batch loss: 0.043
[epoch: 7, i: 99] avg mini-batch loss: 0.012
[epoch: 7, i: 199] avg mini-batch loss: 0.006
[epoch: 7, i: 299] avg mini-batch loss: 0.007
[epoch: 7, i: 399] avg mini-batch loss: 0.013
[epoch: 7, i: 499] avg mini-batch loss: 0.021
[epoch: 7, i: 599] avg mini-batch loss: 0.015
[epoch: 7, i: 699] avg mini-batch loss: 0.033
[epoch: 7, i: 799] avg mini-batch loss: 0.024
[epoch: 7, i: 899] avg mini-batch loss: 0.005
[epoch: 7, i: 999] avg mini-batch loss: 0.012
[epoch: 7, i: 1099] avg mini-batch loss: 0.006
[epoch: 7, i: 1199] avg mini-batch loss: 0.008
[epoch: 7, i: 1299] avg mini-batch loss: 0.005
[epoch: 7, i: 1399] avg mini-batch loss: 0.007
[epoch: 7, i: 1499] avg mini-batch loss: 0.006
[epoch: 7, i: 1599] avg mini-batch loss: 0.008
[epoch: 7, i: 1699] avg mini-batch loss: 0.017
[epoch: 7, i: 1799] avg mini-batch loss: 0.013
[epoch: 7, i: 1899] avg mini-batch loss: 0.002
[epoch: 7, i: 1999] avg mini-batch loss: 0.003
[epoch: 7, i: 2099] avg mini-batch loss: 0.007
[epoch: 7, i: 2199] avg mini-batch loss: 0.004
[epoch: 7, i: 2299] avg mini-batch loss: 0.003
[epoch: 7, i: 2399] avg mini-batch loss: 0.008
[epoch: 7, i: 2499] avg mini-batch loss: 0.007
[epoch: 7, i: 2599] avg mini-batch loss: 0.002
[epoch: 7, i: 2699] avg mini-batch loss: 0.007

[epoch: 7, i: 2799] avg mini-batch loss: 0.012
[epoch: 7, i: 2899] avg mini-batch loss: 0.007
[epoch: 7, i: 2999] avg mini-batch loss: 0.007
[epoch: 7, i: 3099] avg mini-batch loss: 0.002
[epoch: 7, i: 3199] avg mini-batch loss: 0.006
[epoch: 7, i: 3299] avg mini-batch loss: 0.006
[epoch: 7, i: 3399] avg mini-batch loss: 0.011
[epoch: 7, i: 3499] avg mini-batch loss: 0.005
[epoch: 7, i: 3599] avg mini-batch loss: 0.003
[epoch: 7, i: 3699] avg mini-batch loss: 0.003
[epoch: 7, i: 3799] avg mini-batch loss: 0.012
[epoch: 7, i: 3899] avg mini-batch loss: 0.009
[epoch: 7, i: 3999] avg mini-batch loss: 0.002
[epoch: 7, i: 4099] avg mini-batch loss: 0.003
[epoch: 7, i: 4199] avg mini-batch loss: 0.008
[epoch: 7, i: 4299] avg mini-batch loss: 0.002
[epoch: 7, i: 4399] avg mini-batch loss: 0.006
[epoch: 7, i: 4499] avg mini-batch loss: 0.008
[epoch: 7, i: 4599] avg mini-batch loss: 0.003
[epoch: 7, i: 4699] avg mini-batch loss: 0.008
[epoch: 7, i: 4799] avg mini-batch loss: 0.003
[epoch: 7, i: 4899] avg mini-batch loss: 0.005
[epoch: 7, i: 4999] avg mini-batch loss: 0.010
[epoch: 7, i: 5099] avg mini-batch loss: 0.002
[epoch: 7, i: 5199] avg mini-batch loss: 0.001
[epoch: 7, i: 5299] avg mini-batch loss: 0.015
[epoch: 7, i: 5399] avg mini-batch loss: 0.005
[epoch: 7, i: 5499] avg mini-batch loss: 0.006
[epoch: 7, i: 5599] avg mini-batch loss: 0.015
[epoch: 7, i: 5699] avg mini-batch loss: 0.010
[epoch: 7, i: 5799] avg mini-batch loss: 0.013
[epoch: 7, i: 5899] avg mini-batch loss: 0.011
[epoch: 7, i: 5999] avg mini-batch loss: 0.016
[epoch: 7, i: 6099] avg mini-batch loss: 0.023
[epoch: 7, i: 6199] avg mini-batch loss: 0.024
[epoch: 7, i: 6299] avg mini-batch loss: 0.005
[epoch: 7, i: 6399] avg mini-batch loss: 0.003
[epoch: 7, i: 6499] avg mini-batch loss: 0.015
[epoch: 7, i: 6599] avg mini-batch loss: 0.029
[epoch: 7, i: 6699] avg mini-batch loss: 0.006
[epoch: 7, i: 6799] avg mini-batch loss: 0.016
[epoch: 7, i: 6899] avg mini-batch loss: 0.029
[epoch: 7, i: 6999] avg mini-batch loss: 0.034
[epoch: 7, i: 7099] avg mini-batch loss: 0.013
[epoch: 7, i: 7199] avg mini-batch loss: 0.008
[epoch: 7, i: 7299] avg mini-batch loss: 0.028
[epoch: 7, i: 7399] avg mini-batch loss: 0.068
[epoch: 7, i: 7499] avg mini-batch loss: 0.023

[epoch: 7, i: 7599] avg mini-batch loss: 0.043
[epoch: 7, i: 7699] avg mini-batch loss: 0.082
[epoch: 7, i: 7799] avg mini-batch loss: 0.087
[epoch: 7, i: 7899] avg mini-batch loss: 0.022
[epoch: 7, i: 7999] avg mini-batch loss: 0.036
[epoch: 7, i: 8099] avg mini-batch loss: 0.016
[epoch: 7, i: 8199] avg mini-batch loss: 0.034
[epoch: 7, i: 8299] avg mini-batch loss: 0.015
[epoch: 7, i: 8399] avg mini-batch loss: 0.019
[epoch: 7, i: 8499] avg mini-batch loss: 0.020
[epoch: 7, i: 8599] avg mini-batch loss: 0.038
[epoch: 7, i: 8699] avg mini-batch loss: 0.050
[epoch: 7, i: 8799] avg mini-batch loss: 0.020
[epoch: 7, i: 8899] avg mini-batch loss: 0.018
[epoch: 7, i: 8999] avg mini-batch loss: 0.015
[epoch: 7, i: 9099] avg mini-batch loss: 0.027
[epoch: 7, i: 9199] avg mini-batch loss: 0.007
[epoch: 7, i: 9299] avg mini-batch loss: 0.018
[epoch: 7, i: 9399] avg mini-batch loss: 0.044
[epoch: 7, i: 9499] avg mini-batch loss: 0.033
[epoch: 7, i: 9599] avg mini-batch loss: 0.040
[epoch: 7, i: 9699] avg mini-batch loss: 0.008
[epoch: 7, i: 9799] avg mini-batch loss: 0.009
[epoch: 7, i: 9899] avg mini-batch loss: 0.012
[epoch: 7, i: 9999] avg mini-batch loss: 0.012
[epoch: 7, i: 10099] avg mini-batch loss: 0.023
[epoch: 7, i: 10199] avg mini-batch loss: 0.015
[epoch: 7, i: 10299] avg mini-batch loss: 0.031
[epoch: 7, i: 10399] avg mini-batch loss: 0.033
[epoch: 7, i: 10499] avg mini-batch loss: 0.025
[epoch: 7, i: 10599] avg mini-batch loss: 0.046
[epoch: 7, i: 10699] avg mini-batch loss: 0.025
[epoch: 7, i: 10799] avg mini-batch loss: 0.044
[epoch: 7, i: 10899] avg mini-batch loss: 0.041
[epoch: 7, i: 10999] avg mini-batch loss: 0.042
[epoch: 7, i: 11099] avg mini-batch loss: 0.062
[epoch: 7, i: 11199] avg mini-batch loss: 0.087
[epoch: 7, i: 11299] avg mini-batch loss: 0.031
[epoch: 7, i: 11399] avg mini-batch loss: 0.055
[epoch: 7, i: 11499] avg mini-batch loss: 0.017
[epoch: 7, i: 11599] avg mini-batch loss: 0.028
[epoch: 7, i: 11699] avg mini-batch loss: 0.014
[epoch: 7, i: 11799] avg mini-batch loss: 0.022
[epoch: 7, i: 11899] avg mini-batch loss: 0.021
[epoch: 7, i: 11999] avg mini-batch loss: 0.015
[epoch: 7, i: 12099] avg mini-batch loss: 0.018
[epoch: 7, i: 12199] avg mini-batch loss: 0.013
[epoch: 7, i: 12299] avg mini-batch loss: 0.014

[epoch: 7, i: 12399] avg mini-batch loss: 0.006
[epoch: 7, i: 12499] avg mini-batch loss: 0.013
[epoch: 8, i: 99] avg mini-batch loss: 0.023
[epoch: 8, i: 199] avg mini-batch loss: 0.017
[epoch: 8, i: 299] avg mini-batch loss: 0.005
[epoch: 8, i: 399] avg mini-batch loss: 0.008
[epoch: 8, i: 499] avg mini-batch loss: 0.017
[epoch: 8, i: 599] avg mini-batch loss: 0.010
[epoch: 8, i: 699] avg mini-batch loss: 0.015
[epoch: 8, i: 799] avg mini-batch loss: 0.017
[epoch: 8, i: 899] avg mini-batch loss: 0.015
[epoch: 8, i: 999] avg mini-batch loss: 0.010
[epoch: 8, i: 1099] avg mini-batch loss: 0.004
[epoch: 8, i: 1199] avg mini-batch loss: 0.008
[epoch: 8, i: 1299] avg mini-batch loss: 0.008
[epoch: 8, i: 1399] avg mini-batch loss: 0.008
[epoch: 8, i: 1499] avg mini-batch loss: 0.011
[epoch: 8, i: 1599] avg mini-batch loss: 0.011
[epoch: 8, i: 1699] avg mini-batch loss: 0.008
[epoch: 8, i: 1799] avg mini-batch loss: 0.034
[epoch: 8, i: 1899] avg mini-batch loss: 0.036
[epoch: 8, i: 1999] avg mini-batch loss: 0.038
[epoch: 8, i: 2099] avg mini-batch loss: 0.034
[epoch: 8, i: 2199] avg mini-batch loss: 0.026
[epoch: 8, i: 2299] avg mini-batch loss: 0.021
[epoch: 8, i: 2399] avg mini-batch loss: 0.024
[epoch: 8, i: 2499] avg mini-batch loss: 0.020
[epoch: 8, i: 2599] avg mini-batch loss: 0.016
[epoch: 8, i: 2699] avg mini-batch loss: 0.008
[epoch: 8, i: 2799] avg mini-batch loss: 0.018
[epoch: 8, i: 2899] avg mini-batch loss: 0.005
[epoch: 8, i: 2999] avg mini-batch loss: 0.004
[epoch: 8, i: 3099] avg mini-batch loss: 0.006
[epoch: 8, i: 3199] avg mini-batch loss: 0.044
[epoch: 8, i: 3299] avg mini-batch loss: 0.002
[epoch: 8, i: 3399] avg mini-batch loss: 0.012
[epoch: 8, i: 3499] avg mini-batch loss: 0.011
[epoch: 8, i: 3599] avg mini-batch loss: 0.048
[epoch: 8, i: 3699] avg mini-batch loss: 0.074
[epoch: 8, i: 3799] avg mini-batch loss: 0.085
[epoch: 8, i: 3899] avg mini-batch loss: 0.073
[epoch: 8, i: 3999] avg mini-batch loss: 0.024
[epoch: 8, i: 4099] avg mini-batch loss: 0.021
[epoch: 8, i: 4199] avg mini-batch loss: 0.019
[epoch: 8, i: 4299] avg mini-batch loss: 0.006
[epoch: 8, i: 4399] avg mini-batch loss: 0.009
[epoch: 8, i: 4499] avg mini-batch loss: 0.008
[epoch: 8, i: 4599] avg mini-batch loss: 0.055

[epoch: 8, i: 4699] avg mini-batch loss: 0.008
[epoch: 8, i: 4799] avg mini-batch loss: 0.012
[epoch: 8, i: 4899] avg mini-batch loss: 0.009
[epoch: 8, i: 4999] avg mini-batch loss: 0.005
[epoch: 8, i: 5099] avg mini-batch loss: 0.004
[epoch: 8, i: 5199] avg mini-batch loss: 0.008
[epoch: 8, i: 5299] avg mini-batch loss: 0.007
[epoch: 8, i: 5399] avg mini-batch loss: 0.014
[epoch: 8, i: 5499] avg mini-batch loss: 0.059
[epoch: 8, i: 5599] avg mini-batch loss: 0.024
[epoch: 8, i: 5699] avg mini-batch loss: 0.008
[epoch: 8, i: 5799] avg mini-batch loss: 0.014
[epoch: 8, i: 5899] avg mini-batch loss: 0.008
[epoch: 8, i: 5999] avg mini-batch loss: 0.102
[epoch: 8, i: 6099] avg mini-batch loss: 0.073
[epoch: 8, i: 6199] avg mini-batch loss: 0.033
[epoch: 8, i: 6299] avg mini-batch loss: 0.039
[epoch: 8, i: 6399] avg mini-batch loss: 0.012
[epoch: 8, i: 6499] avg mini-batch loss: 0.023
[epoch: 8, i: 6599] avg mini-batch loss: 0.038
[epoch: 8, i: 6699] avg mini-batch loss: 0.027
[epoch: 8, i: 6799] avg mini-batch loss: 0.020
[epoch: 8, i: 6899] avg mini-batch loss: 0.010
[epoch: 8, i: 6999] avg mini-batch loss: 0.011
[epoch: 8, i: 7099] avg mini-batch loss: 0.008
[epoch: 8, i: 7199] avg mini-batch loss: 0.023
[epoch: 8, i: 7299] avg mini-batch loss: 0.007
[epoch: 8, i: 7399] avg mini-batch loss: 0.006
[epoch: 8, i: 7499] avg mini-batch loss: 0.005
[epoch: 8, i: 7599] avg mini-batch loss: 0.003
[epoch: 8, i: 7699] avg mini-batch loss: 0.017
[epoch: 8, i: 7799] avg mini-batch loss: 0.021
[epoch: 8, i: 7899] avg mini-batch loss: 0.006
[epoch: 8, i: 7999] avg mini-batch loss: 0.005
[epoch: 8, i: 8099] avg mini-batch loss: 0.016
[epoch: 8, i: 8199] avg mini-batch loss: 0.033
[epoch: 8, i: 8299] avg mini-batch loss: 0.032
[epoch: 8, i: 8399] avg mini-batch loss: 0.020
[epoch: 8, i: 8499] avg mini-batch loss: 0.004
[epoch: 8, i: 8599] avg mini-batch loss: 0.016
[epoch: 8, i: 8699] avg mini-batch loss: 0.020
[epoch: 8, i: 8799] avg mini-batch loss: 0.011
[epoch: 8, i: 8899] avg mini-batch loss: 0.014
[epoch: 8, i: 8999] avg mini-batch loss: 0.009
[epoch: 8, i: 9099] avg mini-batch loss: 0.014
[epoch: 8, i: 9199] avg mini-batch loss: 0.035
[epoch: 8, i: 9299] avg mini-batch loss: 0.033
[epoch: 8, i: 9399] avg mini-batch loss: 0.021

[epoch: 8, i: 9499] avg mini-batch loss: 0.043
[epoch: 8, i: 9599] avg mini-batch loss: 0.021
[epoch: 8, i: 9699] avg mini-batch loss: 0.036
[epoch: 8, i: 9799] avg mini-batch loss: 0.029
[epoch: 8, i: 9899] avg mini-batch loss: 0.030
[epoch: 8, i: 9999] avg mini-batch loss: 0.006
[epoch: 8, i: 10099] avg mini-batch loss: 0.007
[epoch: 8, i: 10199] avg mini-batch loss: 0.023
[epoch: 8, i: 10299] avg mini-batch loss: 0.036
[epoch: 8, i: 10399] avg mini-batch loss: 0.005
[epoch: 8, i: 10499] avg mini-batch loss: 0.019
[epoch: 8, i: 10599] avg mini-batch loss: 0.017
[epoch: 8, i: 10699] avg mini-batch loss: 0.004
[epoch: 8, i: 10799] avg mini-batch loss: 0.015
[epoch: 8, i: 10899] avg mini-batch loss: 0.013
[epoch: 8, i: 10999] avg mini-batch loss: 0.009
[epoch: 8, i: 11099] avg mini-batch loss: 0.010
[epoch: 8, i: 11199] avg mini-batch loss: 0.013
[epoch: 8, i: 11299] avg mini-batch loss: 0.006
[epoch: 8, i: 11399] avg mini-batch loss: 0.022
[epoch: 8, i: 11499] avg mini-batch loss: 0.012
[epoch: 8, i: 11599] avg mini-batch loss: 0.003
[epoch: 8, i: 11699] avg mini-batch loss: 0.028
[epoch: 8, i: 11799] avg mini-batch loss: 0.038
[epoch: 8, i: 11899] avg mini-batch loss: 0.064
[epoch: 8, i: 11999] avg mini-batch loss: 0.011
[epoch: 8, i: 12099] avg mini-batch loss: 0.034
[epoch: 8, i: 12199] avg mini-batch loss: 0.041
[epoch: 8, i: 12299] avg mini-batch loss: 0.040
[epoch: 8, i: 12399] avg mini-batch loss: 0.026
[epoch: 8, i: 12499] avg mini-batch loss: 0.020
[epoch: 9, i: 99] avg mini-batch loss: 0.014
[epoch: 9, i: 199] avg mini-batch loss: 0.010
[epoch: 9, i: 299] avg mini-batch loss: 0.009
[epoch: 9, i: 399] avg mini-batch loss: 0.006
[epoch: 9, i: 499] avg mini-batch loss: 0.001
[epoch: 9, i: 599] avg mini-batch loss: 0.006
[epoch: 9, i: 699] avg mini-batch loss: 0.003
[epoch: 9, i: 799] avg mini-batch loss: 0.008
[epoch: 9, i: 899] avg mini-batch loss: 0.002
[epoch: 9, i: 999] avg mini-batch loss: 0.016
[epoch: 9, i: 1099] avg mini-batch loss: 0.079
[epoch: 9, i: 1199] avg mini-batch loss: 0.047
[epoch: 9, i: 1299] avg mini-batch loss: 0.020
[epoch: 9, i: 1399] avg mini-batch loss: 0.008
[epoch: 9, i: 1499] avg mini-batch loss: 0.009
[epoch: 9, i: 1599] avg mini-batch loss: 0.023
[epoch: 9, i: 1699] avg mini-batch loss: 0.022

[epoch: 9, i: 1799] avg mini-batch loss: 0.006
[epoch: 9, i: 1899] avg mini-batch loss: 0.036
[epoch: 9, i: 1999] avg mini-batch loss: 0.016
[epoch: 9, i: 2099] avg mini-batch loss: 0.017
[epoch: 9, i: 2199] avg mini-batch loss: 0.006
[epoch: 9, i: 2299] avg mini-batch loss: 0.025
[epoch: 9, i: 2399] avg mini-batch loss: 0.031
[epoch: 9, i: 2499] avg mini-batch loss: 0.011
[epoch: 9, i: 2599] avg mini-batch loss: 0.003
[epoch: 9, i: 2699] avg mini-batch loss: 0.001
[epoch: 9, i: 2799] avg mini-batch loss: 0.008
[epoch: 9, i: 2899] avg mini-batch loss: 0.005
[epoch: 9, i: 2999] avg mini-batch loss: 0.023
[epoch: 9, i: 3099] avg mini-batch loss: 0.003
[epoch: 9, i: 3199] avg mini-batch loss: 0.005
[epoch: 9, i: 3299] avg mini-batch loss: 0.024
[epoch: 9, i: 3399] avg mini-batch loss: 0.015
[epoch: 9, i: 3499] avg mini-batch loss: 0.008
[epoch: 9, i: 3599] avg mini-batch loss: 0.007
[epoch: 9, i: 3699] avg mini-batch loss: 0.005
[epoch: 9, i: 3799] avg mini-batch loss: 0.001
[epoch: 9, i: 3899] avg mini-batch loss: 0.006
[epoch: 9, i: 3999] avg mini-batch loss: 0.008
[epoch: 9, i: 4099] avg mini-batch loss: 0.019
[epoch: 9, i: 4199] avg mini-batch loss: 0.007
[epoch: 9, i: 4299] avg mini-batch loss: 0.001
[epoch: 9, i: 4399] avg mini-batch loss: 0.005
[epoch: 9, i: 4499] avg mini-batch loss: 0.003
[epoch: 9, i: 4599] avg mini-batch loss: 0.003
[epoch: 9, i: 4699] avg mini-batch loss: 0.001
[epoch: 9, i: 4799] avg mini-batch loss: 0.003
[epoch: 9, i: 4899] avg mini-batch loss: 0.003
[epoch: 9, i: 4999] avg mini-batch loss: 0.008
[epoch: 9, i: 5099] avg mini-batch loss: 0.004
[epoch: 9, i: 5199] avg mini-batch loss: 0.001
[epoch: 9, i: 5299] avg mini-batch loss: 0.015
[epoch: 9, i: 5399] avg mini-batch loss: 0.004
[epoch: 9, i: 5499] avg mini-batch loss: 0.047
[epoch: 9, i: 5599] avg mini-batch loss: 0.019
[epoch: 9, i: 5699] avg mini-batch loss: 0.004
[epoch: 9, i: 5799] avg mini-batch loss: 0.004
[epoch: 9, i: 5899] avg mini-batch loss: 0.010
[epoch: 9, i: 5999] avg mini-batch loss: 0.004
[epoch: 9, i: 6099] avg mini-batch loss: 0.002
[epoch: 9, i: 6199] avg mini-batch loss: 0.003
[epoch: 9, i: 6299] avg mini-batch loss: 0.001
[epoch: 9, i: 6399] avg mini-batch loss: 0.004
[epoch: 9, i: 6499] avg mini-batch loss: 0.003

[epoch: 9, i: 6599] avg mini-batch loss: 0.024
[epoch: 9, i: 6699] avg mini-batch loss: 0.018
[epoch: 9, i: 6799] avg mini-batch loss: 0.010
[epoch: 9, i: 6899] avg mini-batch loss: 0.010
[epoch: 9, i: 6999] avg mini-batch loss: 0.010
[epoch: 9, i: 7099] avg mini-batch loss: 0.018
[epoch: 9, i: 7199] avg mini-batch loss: 0.017
[epoch: 9, i: 7299] avg mini-batch loss: 0.014
[epoch: 9, i: 7399] avg mini-batch loss: 0.007
[epoch: 9, i: 7499] avg mini-batch loss: 0.007
[epoch: 9, i: 7599] avg mini-batch loss: 0.044
[epoch: 9, i: 7699] avg mini-batch loss: 0.035
[epoch: 9, i: 7799] avg mini-batch loss: 0.019
[epoch: 9, i: 7899] avg mini-batch loss: 0.014
[epoch: 9, i: 7999] avg mini-batch loss: 0.024
[epoch: 9, i: 8099] avg mini-batch loss: 0.002
[epoch: 9, i: 8199] avg mini-batch loss: 0.008
[epoch: 9, i: 8299] avg mini-batch loss: 0.015
[epoch: 9, i: 8399] avg mini-batch loss: 0.032
[epoch: 9, i: 8499] avg mini-batch loss: 0.017
[epoch: 9, i: 8599] avg mini-batch loss: 0.016
[epoch: 9, i: 8699] avg mini-batch loss: 0.024
[epoch: 9, i: 8799] avg mini-batch loss: 0.020
[epoch: 9, i: 8899] avg mini-batch loss: 0.010
[epoch: 9, i: 8999] avg mini-batch loss: 0.004
[epoch: 9, i: 9099] avg mini-batch loss: 0.024
[epoch: 9, i: 9199] avg mini-batch loss: 0.003
[epoch: 9, i: 9299] avg mini-batch loss: 0.015
[epoch: 9, i: 9399] avg mini-batch loss: 0.009
[epoch: 9, i: 9499] avg mini-batch loss: 0.005
[epoch: 9, i: 9599] avg mini-batch loss: 0.018
[epoch: 9, i: 9699] avg mini-batch loss: 0.023
[epoch: 9, i: 9799] avg mini-batch loss: 0.013
[epoch: 9, i: 9899] avg mini-batch loss: 0.015
[epoch: 9, i: 9999] avg mini-batch loss: 0.038
[epoch: 9, i: 10099] avg mini-batch loss: 0.051
[epoch: 9, i: 10199] avg mini-batch loss: 0.034
[epoch: 9, i: 10299] avg mini-batch loss: 0.034
[epoch: 9, i: 10399] avg mini-batch loss: 0.015
[epoch: 9, i: 10499] avg mini-batch loss: 0.031
[epoch: 9, i: 10599] avg mini-batch loss: 0.039
[epoch: 9, i: 10699] avg mini-batch loss: 0.054
[epoch: 9, i: 10799] avg mini-batch loss: 0.006
[epoch: 9, i: 10899] avg mini-batch loss: 0.013
[epoch: 9, i: 10999] avg mini-batch loss: 0.014
[epoch: 9, i: 11099] avg mini-batch loss: 0.073
[epoch: 9, i: 11199] avg mini-batch loss: 0.019
[epoch: 9, i: 11299] avg mini-batch loss: 0.015

```

[epoch: 9, i: 11399] avg mini-batch loss: 0.004
[epoch: 9, i: 11499] avg mini-batch loss: 0.015
[epoch: 9, i: 11599] avg mini-batch loss: 0.012
[epoch: 9, i: 11699] avg mini-batch loss: 0.010
[epoch: 9, i: 11799] avg mini-batch loss: 0.022
[epoch: 9, i: 11899] avg mini-batch loss: 0.013
[epoch: 9, i: 11999] avg mini-batch loss: 0.016
[epoch: 9, i: 12099] avg mini-batch loss: 0.024
[epoch: 9, i: 12199] avg mini-batch loss: 0.024
[epoch: 9, i: 12299] avg mini-batch loss: 0.008
[epoch: 9, i: 12399] avg mini-batch loss: 0.007
[epoch: 9, i: 12499] avg mini-batch loss: 0.028
Finished Training.

```

```

[23]: correct = 0
total = 0
with torch.no_grad():
    for data in testloader:
        images, labels = data
        images, labels = images.to(device), labels.to(device)
        outputs = net(images)
        _, predicted = torch.max(outputs.data, 1)
        total += labels.size(0)
        correct += (predicted == labels).sum().item()

print('Accuracy of the network on the 10000 test images: %d %%' % (
    100 * correct / total))

```

Accuracy of the network on the 10000 test images: 79 %

```

[24]: class_correct = list(0. for i in range(10))
class_total = list(0. for i in range(10))
with torch.no_grad():
    for data in testloader:
        images, labels = data
        images, labels = images.to(device), labels.to(device)
        outputs = net(images)
        _, predicted = torch.max(outputs, 1)
        c = (predicted == labels).squeeze()
        for i in range(4):
            label = labels[i]
            class_correct[label] += c[i].item()
            class_total[label] += 1

for i in range(10):
    print('Accuracy of %5s : %2d %%' % (
        classes[i], 100 * class_correct[i] / class_total[i]))

```

Accuracy of plane : 83 %

Accuracy of car : 92 %
Accuracy of bird : 70 %
Accuracy of cat : 56 %
Accuracy of deer : 75 %
Accuracy of dog : 71 %
Accuracy of frog : 89 %
Accuracy of horse : 82 %
Accuracy of ship : 88 %
Accuracy of truck : 82 %

[]: