final_11

March 26, 2023

[1]:
```
#!pip install torch torchvision
%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.CIFAR100(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.CIFAR100(root='./data', train=False,
                                        download=True, transform=transform)
testloader = torch.utils.data.DataLoader(testset, batch_size=4,
                                         shuffle=False, num_workers=2)

classes = ('apple', 'aquarium_fish', 'baby', 'bear', 'beaver',
           'bed', 'bee', 'beetle', 'bicycle', 'bottle', 'bowl',
           'boy', 'bridge', 'bus', 'butterfly', 'camel', 'can',
           'castle', 'caterpillar', 'cattle', 'chair',
           'chimpanzee', 'clock', 'cloud', 'cockroach', 'couch',
           'crab', 'crocodile', 'cup', 'dinosaur', 'dolphin',
           'elephant', 'flatfish', 'forest', 'fox', 'girl',
           'hamster', 'house', 'kangaroo', 'keyboard', 'lamp',
           'lawn_mower', 'leopard', 'lion', 'lizard', 'lobster',
           'man', 'maple_tree', 'motorcycle', 'mountain', 'mouse',
           'mushroom', 'oak_tree', 'orange', 'orchid', 'otter',
```

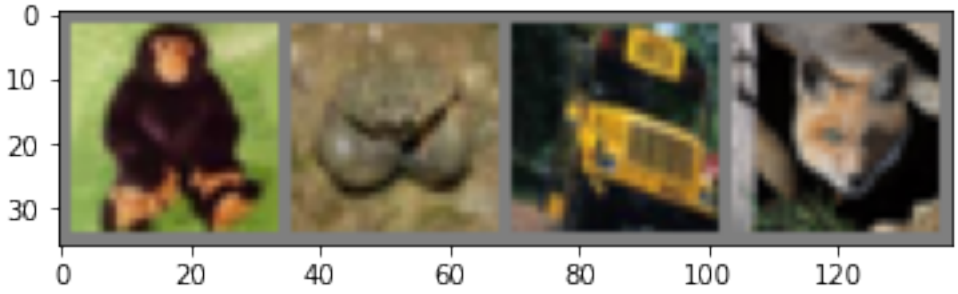
```
           'palm_tree', 'pear', 'pickup_truck', 'pine_tree', 'plain',
           'plate', 'poppy', 'porcupine', 'possum', 'rabbit', 'raccoon',
           'ray', 'road', 'rocket', 'rose', 'sea', 'seal', 'shark',
           'shrew', 'skunk', 'skyscraper', 'snail', 'snake', 'spider',
           'squirrel', 'streetcar', 'sunflower', 'sweet_pepper', 'table',
           'tank', 'telephone', 'television', 'tiger', 'tractor', 'train',
           'trout', 'tulip', 'turtle', 'wardrobe', 'whale', 'willow_tree',
           'wolf', 'woman', 'worm')

#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)))
```




```
chimpanzee  crab   bus   fox
```

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.
```

```
cuda:0
```

Network Definition

[5]:
```
class Net(nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.conv1 = nn.Conv2d(in_channels=3, out_channels=64, kernel_size=3,␣
 ↪padding=1)
        self.bn1 = nn.BatchNorm2d(64)
        self.pool1 = nn.MaxPool2d(kernel_size=2, stride=2)
        self.relu1 = nn.ReLU()
        self.conv2 = nn.Conv2d(64, 128, 3, padding=1)
        self.bn2 = nn.BatchNorm2d(128)
        self.pool2 = nn.MaxPool2d(2, 2)
        self.relu2 = nn.ReLU()
        self.conv3 = nn.Conv2d(128, 256, 3, padding=1)
        self.bn3 = nn.BatchNorm2d(256)
        self.pool3 = nn.MaxPool2d(2, 2)
        self.relu3 = nn.SELU()
        self.fc1 = nn.Linear(256 * 4 * 4, 1024)
        self.bn4 = nn.BatchNorm1d(1024)
        self.relu4 = nn.SELU()
        self.fc2 = nn.Linear(1024, 100)

    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()      # Create the network instance.
net.to(device)   # Move the network parameters to the specified device.
```

[5]:
```
Net(
  (conv1): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (pool1): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
```

```
  (relu1): ReLU()
  (conv2): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (pool2): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (relu2): ReLU()
  (conv3): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (pool3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (relu3): SELU()
  (fc1): Linear(in_features=4096, out_features=1024, bias=True)
  (bn4): BatchNorm1d(1024, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (relu4): SELU()
  (fc2): Linear(in_features=1024, out_features=100, bias=True)
)
```

Optimizer and Loss Function

[6]:
```python
# 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=0.0001, momentum=0.9)
opt = optim.Adam(net.parameters(), lr=0.0001)
```

Training Procedure

[7]:
```python
import sys
from tqdm.notebook import tqdm

avg_losses = []   # Avg. losses.
epochs = 10       # Total epochs.
print_freq = 500  # 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(tqdm(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),flush=True)
            sys.stdout.flush()
            avg_losses.append(avg_loss)
            running_loss = 0.0

print('Finished Training.')
```

```
  0%|          | 0/12500 [00:00<?, ?it/s]
[epoch: 0, i:   499] avg mini-batch loss: 4.387
[epoch: 0, i:   999] avg mini-batch loss: 4.033
[epoch: 0, i:  1499] avg mini-batch loss: 3.796
[epoch: 0, i:  1999] avg mini-batch loss: 3.606
[epoch: 0, i:  2499] avg mini-batch loss: 3.541
[epoch: 0, i:  2999] avg mini-batch loss: 3.467
[epoch: 0, i:  3499] avg mini-batch loss: 3.386
[epoch: 0, i:  3999] avg mini-batch loss: 3.335
[epoch: 0, i:  4499] avg mini-batch loss: 3.167
[epoch: 0, i:  4999] avg mini-batch loss: 3.101
[epoch: 0, i:  5499] avg mini-batch loss: 3.030
[epoch: 0, i:  5999] avg mini-batch loss: 2.934
[epoch: 0, i:  6499] avg mini-batch loss: 2.886
[epoch: 0, i:  6999] avg mini-batch loss: 2.986
[epoch: 0, i:  7499] avg mini-batch loss: 2.942
[epoch: 0, i:  7999] avg mini-batch loss: 2.872
[epoch: 0, i:  8499] avg mini-batch loss: 2.874
[epoch: 0, i:  8999] avg mini-batch loss: 2.869
[epoch: 0, i:  9499] avg mini-batch loss: 2.760
[epoch: 0, i:  9999] avg mini-batch loss: 2.794
[epoch: 0, i: 10499] avg mini-batch loss: 2.748
[epoch: 0, i: 10999] avg mini-batch loss: 2.638
```

```
[epoch: 0, i: 11499] avg mini-batch loss: 2.629
[epoch: 0, i: 11999] avg mini-batch loss: 2.695
[epoch: 0, i: 12499] avg mini-batch loss: 2.671
```

```
  0%|          | 0/12500 [00:00<?, ?it/s]
```

```
[epoch: 1, i:   499] avg mini-batch loss: 2.510
[epoch: 1, i:   999] avg mini-batch loss: 2.459
[epoch: 1, i:  1499] avg mini-batch loss: 2.427
[epoch: 1, i:  1999] avg mini-batch loss: 2.480
[epoch: 1, i:  2499] avg mini-batch loss: 2.447
[epoch: 1, i:  2999] avg mini-batch loss: 2.401
[epoch: 1, i:  3499] avg mini-batch loss: 2.450
[epoch: 1, i:  3999] avg mini-batch loss: 2.431
[epoch: 1, i:  4499] avg mini-batch loss: 2.331
[epoch: 1, i:  4999] avg mini-batch loss: 2.422
[epoch: 1, i:  5499] avg mini-batch loss: 2.383
[epoch: 1, i:  5999] avg mini-batch loss: 2.389
[epoch: 1, i:  6499] avg mini-batch loss: 2.387
[epoch: 1, i:  6999] avg mini-batch loss: 2.279
[epoch: 1, i:  7499] avg mini-batch loss: 2.322
[epoch: 1, i:  7999] avg mini-batch loss: 2.363
[epoch: 1, i:  8499] avg mini-batch loss: 2.278
[epoch: 1, i:  8999] avg mini-batch loss: 2.303
[epoch: 1, i:  9499] avg mini-batch loss: 2.274
[epoch: 1, i:  9999] avg mini-batch loss: 2.288
[epoch: 1, i: 10499] avg mini-batch loss: 2.266
[epoch: 1, i: 10999] avg mini-batch loss: 2.279
[epoch: 1, i: 11499] avg mini-batch loss: 2.300
[epoch: 1, i: 11999] avg mini-batch loss: 2.282
[epoch: 1, i: 12499] avg mini-batch loss: 2.312
```

```
  0%|          | 0/12500 [00:00<?, ?it/s]
```

```
[epoch: 2, i:   499] avg mini-batch loss: 1.868
[epoch: 2, i:   999] avg mini-batch loss: 1.952
[epoch: 2, i:  1499] avg mini-batch loss: 1.981
[epoch: 2, i:  1999] avg mini-batch loss: 2.006
[epoch: 2, i:  2499] avg mini-batch loss: 1.958
[epoch: 2, i:  2999] avg mini-batch loss: 2.004
[epoch: 2, i:  3499] avg mini-batch loss: 2.093
[epoch: 2, i:  3999] avg mini-batch loss: 2.069
[epoch: 2, i:  4499] avg mini-batch loss: 1.873
[epoch: 2, i:  4999] avg mini-batch loss: 1.944
[epoch: 2, i:  5499] avg mini-batch loss: 1.958
[epoch: 2, i:  5999] avg mini-batch loss: 2.058
[epoch: 2, i:  6499] avg mini-batch loss: 2.034
[epoch: 2, i:  6999] avg mini-batch loss: 2.035
[epoch: 2, i:  7499] avg mini-batch loss: 1.979
[epoch: 2, i:  7999] avg mini-batch loss: 1.894
```

```
[epoch: 2, i:  8499] avg mini-batch loss: 1.953
[epoch: 2, i:  8999] avg mini-batch loss: 2.015
[epoch: 2, i:  9499] avg mini-batch loss: 1.983
[epoch: 2, i:  9999] avg mini-batch loss: 1.974
[epoch: 2, i: 10499] avg mini-batch loss: 1.964
[epoch: 2, i: 10999] avg mini-batch loss: 2.052
[epoch: 2, i: 11499] avg mini-batch loss: 1.931
[epoch: 2, i: 11999] avg mini-batch loss: 2.006
[epoch: 2, i: 12499] avg mini-batch loss: 2.021
  0%|          | 0/12500 [00:00<?, ?it/s]
[epoch: 3, i:   499] avg mini-batch loss: 1.603
[epoch: 3, i:   999] avg mini-batch loss: 1.621
[epoch: 3, i:  1499] avg mini-batch loss: 1.577
[epoch: 3, i:  1999] avg mini-batch loss: 1.636
[epoch: 3, i:  2499] avg mini-batch loss: 1.622
[epoch: 3, i:  2999] avg mini-batch loss: 1.646
[epoch: 3, i:  3499] avg mini-batch loss: 1.619
[epoch: 3, i:  3999] avg mini-batch loss: 1.660
[epoch: 3, i:  4499] avg mini-batch loss: 1.730
[epoch: 3, i:  4999] avg mini-batch loss: 1.674
[epoch: 3, i:  5499] avg mini-batch loss: 1.679
[epoch: 3, i:  5999] avg mini-batch loss: 1.692
[epoch: 3, i:  6499] avg mini-batch loss: 1.672
[epoch: 3, i:  6999] avg mini-batch loss: 1.649
[epoch: 3, i:  7499] avg mini-batch loss: 1.673
[epoch: 3, i:  7999] avg mini-batch loss: 1.642
[epoch: 3, i:  8499] avg mini-batch loss: 1.635
[epoch: 3, i:  8999] avg mini-batch loss: 1.657
[epoch: 3, i:  9499] avg mini-batch loss: 1.691
[epoch: 3, i:  9999] avg mini-batch loss: 1.716
[epoch: 3, i: 10499] avg mini-batch loss: 1.637
[epoch: 3, i: 10999] avg mini-batch loss: 1.672
[epoch: 3, i: 11499] avg mini-batch loss: 1.662
[epoch: 3, i: 11999] avg mini-batch loss: 1.720
[epoch: 3, i: 12499] avg mini-batch loss: 1.633
  0%|          | 0/12500 [00:00<?, ?it/s]
[epoch: 4, i:   499] avg mini-batch loss: 1.204
[epoch: 4, i:   999] avg mini-batch loss: 1.231
[epoch: 4, i:  1499] avg mini-batch loss: 1.271
[epoch: 4, i:  1999] avg mini-batch loss: 1.291
[epoch: 4, i:  2499] avg mini-batch loss: 1.276
[epoch: 4, i:  2999] avg mini-batch loss: 1.257
[epoch: 4, i:  3499] avg mini-batch loss: 1.281
[epoch: 4, i:  3999] avg mini-batch loss: 1.302
[epoch: 4, i:  4499] avg mini-batch loss: 1.302
[epoch: 4, i:  4999] avg mini-batch loss: 1.310
```

```
[epoch: 4, i:  5499] avg mini-batch loss: 1.352
[epoch: 4, i:  5999] avg mini-batch loss: 1.271
[epoch: 4, i:  6499] avg mini-batch loss: 1.325
[epoch: 4, i:  6999] avg mini-batch loss: 1.253
[epoch: 4, i:  7499] avg mini-batch loss: 1.369
[epoch: 4, i:  7999] avg mini-batch loss: 1.413
[epoch: 4, i:  8499] avg mini-batch loss: 1.346
[epoch: 4, i:  8999] avg mini-batch loss: 1.310
[epoch: 4, i:  9499] avg mini-batch loss: 1.438
[epoch: 4, i:  9999] avg mini-batch loss: 1.411
[epoch: 4, i: 10499] avg mini-batch loss: 1.353
[epoch: 4, i: 10999] avg mini-batch loss: 1.371
[epoch: 4, i: 11499] avg mini-batch loss: 1.427
[epoch: 4, i: 11999] avg mini-batch loss: 1.375
[epoch: 4, i: 12499] avg mini-batch loss: 1.339

  0%|          | 0/12500 [00:00<?, ?it/s]

[epoch: 5, i:   499] avg mini-batch loss: 0.884
[epoch: 5, i:   999] avg mini-batch loss: 0.895
[epoch: 5, i:  1499] avg mini-batch loss: 0.897
[epoch: 5, i:  1999] avg mini-batch loss: 0.903
[epoch: 5, i:  2499] avg mini-batch loss: 0.918
[epoch: 5, i:  2999] avg mini-batch loss: 0.927
[epoch: 5, i:  3499] avg mini-batch loss: 0.968
[epoch: 5, i:  3999] avg mini-batch loss: 1.007
[epoch: 5, i:  4499] avg mini-batch loss: 1.048
[epoch: 5, i:  4999] avg mini-batch loss: 1.049
[epoch: 5, i:  5499] avg mini-batch loss: 1.029
[epoch: 5, i:  5999] avg mini-batch loss: 0.993
[epoch: 5, i:  6499] avg mini-batch loss: 0.984
[epoch: 5, i:  6999] avg mini-batch loss: 0.976
[epoch: 5, i:  7499] avg mini-batch loss: 0.967
[epoch: 5, i:  7999] avg mini-batch loss: 1.007
[epoch: 5, i:  8499] avg mini-batch loss: 1.020
[epoch: 5, i:  8999] avg mini-batch loss: 1.030
[epoch: 5, i:  9499] avg mini-batch loss: 1.015
[epoch: 5, i:  9999] avg mini-batch loss: 1.000
[epoch: 5, i: 10499] avg mini-batch loss: 1.091
[epoch: 5, i: 10999] avg mini-batch loss: 0.979
[epoch: 5, i: 11499] avg mini-batch loss: 0.995
[epoch: 5, i: 11999] avg mini-batch loss: 1.046
[epoch: 5, i: 12499] avg mini-batch loss: 1.059

  0%|          | 0/12500 [00:00<?, ?it/s]

[epoch: 6, i:   499] avg mini-batch loss: 0.498
[epoch: 6, i:   999] avg mini-batch loss: 0.552
[epoch: 6, i:  1499] avg mini-batch loss: 0.603
[epoch: 6, i:  1999] avg mini-batch loss: 0.576
```

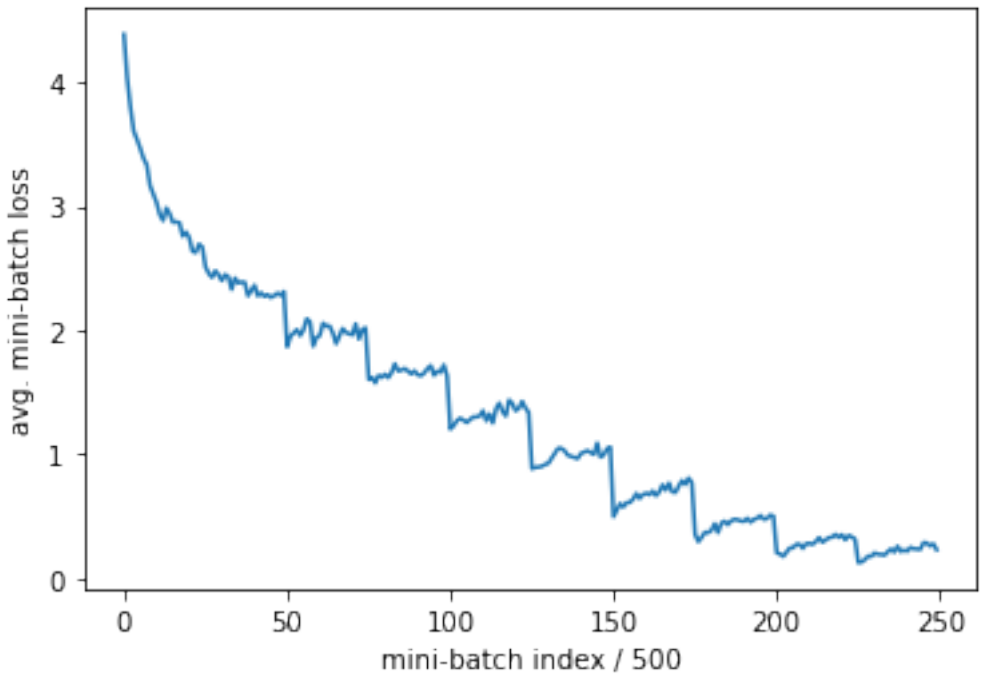
```
[epoch: 6, i:  2499] avg mini-batch loss: 0.611
[epoch: 6, i:  2999] avg mini-batch loss: 0.613
[epoch: 6, i:  3499] avg mini-batch loss: 0.639
[epoch: 6, i:  3999] avg mini-batch loss: 0.685
[epoch: 6, i:  4499] avg mini-batch loss: 0.644
[epoch: 6, i:  4999] avg mini-batch loss: 0.676
[epoch: 6, i:  5499] avg mini-batch loss: 0.686
[epoch: 6, i:  5999] avg mini-batch loss: 0.675
[epoch: 6, i:  6499] avg mini-batch loss: 0.705
[epoch: 6, i:  6999] avg mini-batch loss: 0.667
[epoch: 6, i:  7499] avg mini-batch loss: 0.696
[epoch: 6, i:  7999] avg mini-batch loss: 0.754
[epoch: 6, i:  8499] avg mini-batch loss: 0.712
[epoch: 6, i:  8999] avg mini-batch loss: 0.768
[epoch: 6, i:  9499] avg mini-batch loss: 0.697
[epoch: 6, i:  9999] avg mini-batch loss: 0.694
[epoch: 6, i: 10499] avg mini-batch loss: 0.745
[epoch: 6, i: 10999] avg mini-batch loss: 0.789
[epoch: 6, i: 11499] avg mini-batch loss: 0.757
[epoch: 6, i: 11999] avg mini-batch loss: 0.810
[epoch: 6, i: 12499] avg mini-batch loss: 0.771
  0%|          | 0/12500 [00:00<?, ?it/s]
[epoch: 7, i:   499] avg mini-batch loss: 0.354
[epoch: 7, i:   999] avg mini-batch loss: 0.296
[epoch: 7, i:  1499] avg mini-batch loss: 0.331
[epoch: 7, i:  1999] avg mini-batch loss: 0.372
[epoch: 7, i:  2499] avg mini-batch loss: 0.369
[epoch: 7, i:  2999] avg mini-batch loss: 0.388
[epoch: 7, i:  3499] avg mini-batch loss: 0.440
[epoch: 7, i:  3999] avg mini-batch loss: 0.369
[epoch: 7, i:  4499] avg mini-batch loss: 0.445
[epoch: 7, i:  4999] avg mini-batch loss: 0.461
[epoch: 7, i:  5499] avg mini-batch loss: 0.433
[epoch: 7, i:  5999] avg mini-batch loss: 0.464
[epoch: 7, i:  6499] avg mini-batch loss: 0.476
[epoch: 7, i:  6999] avg mini-batch loss: 0.473
[epoch: 7, i:  7499] avg mini-batch loss: 0.463
[epoch: 7, i:  7999] avg mini-batch loss: 0.460
[epoch: 7, i:  8499] avg mini-batch loss: 0.485
[epoch: 7, i:  8999] avg mini-batch loss: 0.454
[epoch: 7, i:  9499] avg mini-batch loss: 0.475
[epoch: 7, i:  9999] avg mini-batch loss: 0.485
[epoch: 7, i: 10499] avg mini-batch loss: 0.507
[epoch: 7, i: 10999] avg mini-batch loss: 0.477
[epoch: 7, i: 11499] avg mini-batch loss: 0.482
[epoch: 7, i: 11999] avg mini-batch loss: 0.510
[epoch: 7, i: 12499] avg mini-batch loss: 0.501
```

```
  0%|          | 0/12500 [00:00<?, ?it/s]
[epoch: 8, i:   499] avg mini-batch loss: 0.206
[epoch: 8, i:   999] avg mini-batch loss: 0.197
[epoch: 8, i:  1499] avg mini-batch loss: 0.178
[epoch: 8, i:  1999] avg mini-batch loss: 0.212
[epoch: 8, i:  2499] avg mini-batch loss: 0.245
[epoch: 8, i:  2999] avg mini-batch loss: 0.247
[epoch: 8, i:  3499] avg mini-batch loss: 0.274
[epoch: 8, i:  3999] avg mini-batch loss: 0.276
[epoch: 8, i:  4499] avg mini-batch loss: 0.242
[epoch: 8, i:  4999] avg mini-batch loss: 0.280
[epoch: 8, i:  5499] avg mini-batch loss: 0.282
[epoch: 8, i:  5999] avg mini-batch loss: 0.275
[epoch: 8, i:  6499] avg mini-batch loss: 0.297
[epoch: 8, i:  6999] avg mini-batch loss: 0.329
[epoch: 8, i:  7499] avg mini-batch loss: 0.292
[epoch: 8, i:  7999] avg mini-batch loss: 0.310
[epoch: 8, i:  8499] avg mini-batch loss: 0.329
[epoch: 8, i:  8999] avg mini-batch loss: 0.332
[epoch: 8, i:  9499] avg mini-batch loss: 0.354
[epoch: 8, i:  9999] avg mini-batch loss: 0.331
[epoch: 8, i: 10499] avg mini-batch loss: 0.352
[epoch: 8, i: 10999] avg mini-batch loss: 0.311
[epoch: 8, i: 11499] avg mini-batch loss: 0.346
[epoch: 8, i: 11999] avg mini-batch loss: 0.334
[epoch: 8, i: 12499] avg mini-batch loss: 0.317
  0%|          | 0/12500 [00:00<?, ?it/s]
[epoch: 9, i:   499] avg mini-batch loss: 0.128
[epoch: 9, i:   999] avg mini-batch loss: 0.132
[epoch: 9, i:  1499] avg mini-batch loss: 0.151
[epoch: 9, i:  1999] avg mini-batch loss: 0.180
[epoch: 9, i:  2499] avg mini-batch loss: 0.178
[epoch: 9, i:  2999] avg mini-batch loss: 0.205
[epoch: 9, i:  3499] avg mini-batch loss: 0.194
[epoch: 9, i:  3999] avg mini-batch loss: 0.193
[epoch: 9, i:  4499] avg mini-batch loss: 0.187
[epoch: 9, i:  4999] avg mini-batch loss: 0.214
[epoch: 9, i:  5499] avg mini-batch loss: 0.236
[epoch: 9, i:  5999] avg mini-batch loss: 0.212
[epoch: 9, i:  6499] avg mini-batch loss: 0.259
[epoch: 9, i:  6999] avg mini-batch loss: 0.217
[epoch: 9, i:  7499] avg mini-batch loss: 0.228
[epoch: 9, i:  7999] avg mini-batch loss: 0.219
[epoch: 9, i:  8499] avg mini-batch loss: 0.250
[epoch: 9, i:  8999] avg mini-batch loss: 0.239
[epoch: 9, i:  9499] avg mini-batch loss: 0.237
```

```
[epoch: 9, i:  9999] avg mini-batch loss: 0.235
[epoch: 9, i: 10499] avg mini-batch loss: 0.288
[epoch: 9, i: 10999] avg mini-batch loss: 0.285
[epoch: 9, i: 11499] avg mini-batch loss: 0.262
[epoch: 9, i: 11999] avg mini-batch loss: 0.278
[epoch: 9, i: 12499] avg mini-batch loss: 0.231
Finished Training.
```

Training Loss Curve

[8]:
```
plt.plot(avg_losses)
plt.xlabel('mini-batch index / {}'.format(print_freq))
plt.ylabel('avg. mini-batch loss')
plt.show()
```

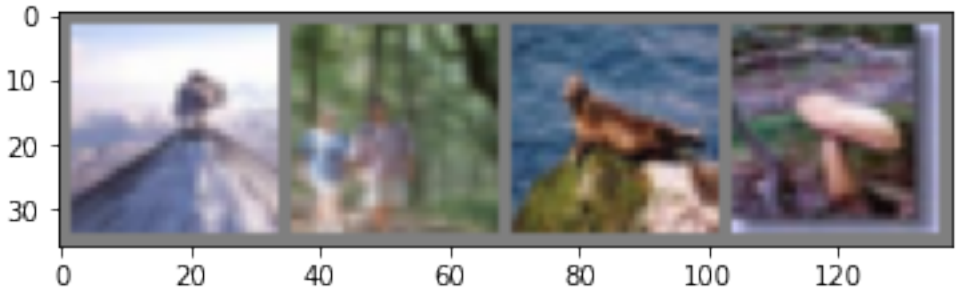


Evaluate on Test Dataset

[9]:
```
# 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:  mountain forest  seal mushroom
Predicted:   road elephant dolphin apple
```

[10]:
```
# 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: 43 %
```

[11]:
```
# Get test accuracy for each class.
class_correct = [0] * len(classes)
class_total = [0] * len(classes)
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(len(labels)):
            label = labels[i]
            class_correct[label] += c[i].item()
```

```
            class_total[label] += 1

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

```
Accuracy of apple : 68 %
Accuracy of aquarium_fish : 57 %
Accuracy of  baby : 24 %
Accuracy of  bear : 25 %
Accuracy of beaver : 26 %
Accuracy of   bed : 45 %
Accuracy of   bee : 55 %
Accuracy of beetle : 57 %
Accuracy of bicycle : 60 %
Accuracy of bottle : 61 %
Accuracy of  bowl : 26 %
Accuracy of   boy : 24 %
Accuracy of bridge : 46 %
Accuracy of   bus : 47 %
Accuracy of butterfly : 41 %
Accuracy of camel : 34 %
Accuracy of   can : 39 %
Accuracy of castle : 54 %
Accuracy of caterpillar : 51 %
Accuracy of cattle : 36 %
Accuracy of chair : 73 %
Accuracy of chimpanzee : 52 %
Accuracy of clock : 38 %
Accuracy of cloud : 65 %
Accuracy of cockroach : 76 %
Accuracy of couch : 33 %
Accuracy of  crab : 33 %
Accuracy of crocodile : 15 %
Accuracy of   cup : 70 %
Accuracy of dinosaur : 36 %
Accuracy of dolphin : 50 %
Accuracy of elephant : 41 %
Accuracy of flatfish : 32 %
Accuracy of forest : 50 %
Accuracy of   fox : 32 %
Accuracy of  girl : 30 %
Accuracy of hamster : 50 %
Accuracy of house : 43 %
Accuracy of kangaroo : 34 %
Accuracy of keyboard : 47 %
Accuracy of  lamp : 29 %
```

```
Accuracy of lawn_mower : 57 %
Accuracy of leopard : 44 %
Accuracy of  lion : 41 %
Accuracy of lizard : 25 %
Accuracy of lobster : 32 %
Accuracy of   man : 27 %
Accuracy of maple_tree : 46 %
Accuracy of motorcycle : 73 %
Accuracy of mountain : 44 %
Accuracy of mouse : 17 %
Accuracy of mushroom : 43 %
Accuracy of oak_tree : 55 %
Accuracy of orange : 65 %
Accuracy of orchid : 60 %
Accuracy of otter : 13 %
Accuracy of palm_tree : 53 %
Accuracy of  pear : 48 %
Accuracy of pickup_truck : 57 %
Accuracy of pine_tree : 45 %
Accuracy of plain : 81 %
Accuracy of plate : 55 %
Accuracy of poppy : 42 %
Accuracy of porcupine : 45 %
Accuracy of possum : 16 %
Accuracy of rabbit : 22 %
Accuracy of raccoon : 40 %
Accuracy of   ray : 44 %
Accuracy of  road : 62 %
Accuracy of rocket : 63 %
Accuracy of  rose : 13 %
Accuracy of   sea : 46 %
Accuracy of  seal : 17 %
Accuracy of shark : 33 %
Accuracy of shrew : 19 %
Accuracy of skunk : 68 %
Accuracy of skyscraper : 72 %
Accuracy of snail : 32 %
Accuracy of snake : 12 %
Accuracy of spider : 47 %
Accuracy of squirrel : 27 %
Accuracy of streetcar : 44 %
Accuracy of sunflower : 69 %
Accuracy of sweet_pepper : 40 %
Accuracy of table : 45 %
Accuracy of  tank : 62 %
Accuracy of telephone : 46 %
Accuracy of television : 46 %
Accuracy of tiger : 36 %
```

```
Accuracy of tractor : 45 %
Accuracy of train : 32 %
Accuracy of trout : 51 %
Accuracy of tulip : 59 %
Accuracy of turtle : 25 %
Accuracy of wardrobe : 80 %
Accuracy of whale : 36 %
Accuracy of willow_tree : 40 %
Accuracy of  wolf : 54 %
Accuracy of woman : 20 %
Accuracy of  worm : 42 %
```

[12]:
```
# One of the changes I made was that I added another layer in the network
# that takes the output from the second convolutional layer, applies ReLU␣
 ↪activation,
# and passes it through a 2x2 AvgPool layer to capture more relationships.
# I also reduced the learning rate of the optimizer to half (0.0005) which
# which allowed the optimizer to take smaller steps towards the minimum of
# the loss function which might allow for a better chance of finding the global␣
 ↪min of loss.
```