dog_app

September 23, 2019

1 Convolutional Neural Networks

1.1 Project: Write an Algorithm for a Dog Identification App

In this notebook, some template code has already been provided for you, and you will need to implement additional functionality to successfully complete this project. You will not need to modify the included code beyond what is requested. Sections that begin with **'(IMPLEMENTATION)'** in the header indicate that the following block of code will require additional functionality which you must provide. Instructions will be provided for each section, and the specifics of the implementation are marked in the code block with a 'TODO' statement. Please be sure to read the instructions carefully!

> **Note**: Once you have completed all of the code implementations, you need to finalize your work by exporting the Jupyter Notebook as an HTML document. Before exporting the notebook to html, all of the code cells need to have been run so that reviewers can see the final implementation and output. You can then export the notebook by using the menu above and navigating to **File -> Download as -> HTML (.html)**. Include the finished document along with this notebook as your submission.

In addition to implementing code, there will be questions that you must answer which relate to the project and your implementation. Each section where you will answer a question is preceded by a **'Question X'** header. Carefully read each question and provide thorough answers in the following text boxes that begin with **'Answer:'**. Your project submission will be evaluated based on your answers to each of the questions and the implementation you provide.

> **Note:** Code and Markdown cells can be executed using the **Shift + Enter** keyboard shortcut. Markdown cells can be edited by double-clicking the cell to enter edit mode.

The rubric contains *optional* "Stand Out Suggestions" for enhancing the project beyond the minimum requirements. If you decide to pursue the "Stand Out Suggestions", you should include the code in this Jupyter notebook.

Step 0: Import Datasets

Make sure that you've downloaded the required human and dog datasets:

Note: if you are using the Udacity workspace, you *DO NOT* need to re-download these - they can be found in the `/data` folder as noted in the cell below.

- Download the dog dataset. Unzip the folder and place it in this project's home directory, at the location /dog_images.
- Download the human dataset. Unzip the folder and place it in the home directory, at location /lfw.

Note: If you are using a Windows machine, you are encouraged to use 7zip to extract the folder.

In the code cell below, we save the file paths for both the human (LFW) dataset and dog dataset in the numpy arrays human_files and dog_files.

```
In [3]: import numpy as np
        from glob import glob

        # load filenames for human and dog images
        human_files = np.array(glob("/data/lfw/*/*"))
        dog_files = np.array(glob("/data/dog_images/*/*/*"))

        # print number of images in each dataset
        print('There are %d total human images.' % len(human_files))
        print('There are %d total dog images.' % len(dog_files))
```

```
There are 13233 total human images.
There are 8351 total dog images.
```

Step 1: Detect Humans

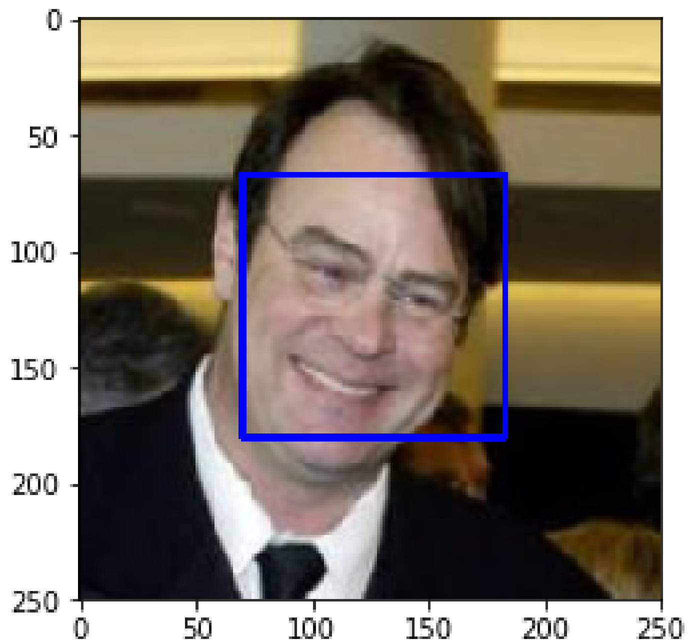
In this section, we use OpenCV's implementation of Haar feature-based cascade classifiers to detect human faces in images.

OpenCV provides many pre-trained face detectors, stored as XML files on github. We have downloaded one of these detectors and stored it in the haarcascades directory. In the next code cell, we demonstrate how to use this detector to find human faces in a sample image.

```
In [4]: import cv2
        import matplotlib.pyplot as plt
        %matplotlib inline

        # extract pre-trained face detector
        face_cascade = cv2.CascadeClassifier('haarcascades/haarcascade_frontalface_alt.xml')

        # load color (BGR) image
        img = cv2.imread(human_files[0])
        # convert BGR image to grayscale
        gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

        # find faces in image
        faces = face_cascade.detectMultiScale(gray)

        # print number of faces detected in the image
        print('Number of faces detected:', len(faces))
```

```python
# get bounding box for each detected face
for (x,y,w,h) in faces:
    # add bounding box to color image
    cv2.rectangle(img,(x,y),(x+w,y+h),(255,0,0),2)

# convert BGR image to RGB for plotting
cv_rgb = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

# display the image, along with bounding box
plt.imshow(cv_rgb)
plt.show()
```

```
Number of faces detected: 1
```



Before using any of the face detectors, it is standard procedure to convert the images to grayscale. The `detectMultiScale` function executes the classifier stored in `face_cascade` and takes the grayscale image as a parameter.

In the above code, `faces` is a numpy array of detected faces, where each row corresponds to a detected face. Each detected face is a 1D array with four entries that specifies the bounding box of the detected face. The first two entries in the array (extracted in the above code as `x` and `y`) specify the horizontal and vertical positions of the top left corner of the bounding box. The last two entries in the array (extracted here as `w` and `h`) specify the width and height of the box.

1.1.1 Write a Human Face Detector

We can use this procedure to write a function that returns `True` if a human face is detected in an image and `False` otherwise. This function, aptly named `face_detector`, takes a string-valued file path to an image as input and appears in the code block below.

```
In [5]: # returns "True" if face is detected in image stored at img_path
        def face_detector(img_path):
            img = cv2.imread(img_path)
            gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
            faces = face_cascade.detectMultiScale(gray)
            return len(faces) > 0
```

1.1.2 (IMPLEMENTATION) Assess the Human Face Detector

Question 1: Use the code cell below to test the performance of the `face_detector` function.
- What percentage of the first 100 images in `human_files` have a detected human face?
- What percentage of the first 100 images in `dog_files` have a detected human face?

Ideally, we would like 100% of human images with a detected face and 0% of dog images with a detected face. You will see that our algorithm falls short of this goal, but still gives acceptable performance. We extract the file paths for the first 100 images from each of the datasets and store them in the numpy arrays `human_files_short` and `dog_files_short`.

Answer: (You can print out your results and/or write your percentages in this cell)

```
In [6]: from tqdm import tqdm

        human_files_short = human_files[:100]
        dog_files_short = dog_files[:100]

        #-#-# Do NOT modify the code above this line. #-#-#

        ## TODO: Test the performance of the face_detector algorithm
        ## on the images in human_files_short and dog_files_short.
        human_face_count = 0
        dog_face_count = 0
        for human_file in human_files_short:
            if face_detector(human_file):
                human_face_count += 1
        print('Percentage of human face detection', human_face_count / len(human_files_short))

        for dog_file in dog_files_short:
            if face_detector(dog_file):
                dog_face_count += 1
        print('Percentage of dog face detection', dog_face_count / len(dog_files_short))
```

```
Percentage of human face detection 0.98
Percentage of dog face detection 0.17
```

We suggest the face detector from OpenCV as a potential way to detect human images in your algorithm, but you are free to explore other approaches, especially approaches that make use of deep learning :). Please use the code cell below to design and test your own face detection algorithm. If you decide to pursue this *optional* task, report performance on `human_files_short` and `dog_files_short`.

```
In [5]: ### (Optional)
        ### TODO: Test performance of anotherface detection algorithm.
        ### Feel free to use as many code cells as needed.
```

Step 2: Detect Dogs
In this section, we use a pre-trained model to detect dogs in images.

1.1.3 Obtain Pre-trained VGG-16 Model

The code cell below downloads the VGG-16 model, along with weights that have been trained on ImageNet, a very large, very popular dataset used for image classification and other vision tasks. ImageNet contains over 10 million URLs, each linking to an image containing an object from one of 1000 categories.

```
In [2]: import torch
        import torchvision.models as models

        # define VGG16 model
        VGG16 = models.vgg16(pretrained=True)

        # check if CUDA is available
        use_cuda = torch.cuda.is_available()

        # move model to GPU if CUDA is available
        if use_cuda:
            VGG16 = VGG16.cuda()
```

```
Downloading: "https://download.pytorch.org/models/vgg16-397923af.pth" to /root/.torch/models/vgg
100%|| 553433881/553433881 [00:08<00:00, 62351022.05it/s]
```

Given an image, this pre-trained VGG-16 model returns a prediction (derived from the 1000 possible categories in ImageNet) for the object that is contained in the image.

1.1.4 (IMPLEMENTATION) Making Predictions with a Pre-trained Model

In the next code cell, you will write a function that accepts a path to an image (such as `'dogImages/train/001.Affenpinscher/Affenpinscher_00001.jpg'`) as input and returns the index corresponding to the ImageNet class that is predicted by the pre-trained VGG-16 model. The output should always be an integer between 0 and 999, inclusive.

Before writing the function, make sure that you take the time to learn how to appropriately pre-process tensors for pre-trained models in the PyTorch documentation.

In [7]:
```python
from PIL import Image
import torchvision.transforms as transforms
from torchvision import datasets
from torch.autograd import Variable

def VGG16_predict(img_path):
    '''
    Use pre-trained VGG-16 model to obtain index corresponding to
    predicted ImageNet class for image at specified path

    Args:
        img_path: path to an image

    Returns:
        Index corresponding to VGG-16 model's prediction
    '''

    ## TODO: Complete the function.
    ## Load and pre-process an image from the given img_path
    ## Return the *index* of the predicted class for that image



    normalize = transforms.Normalize(mean=[0.485, 0.456, 0.406],
                                     std=[0.229, 0.224, 0.225])

    data_transform = transforms.Compose([transforms.Scale(256),
                                         transforms.CenterCrop(224),
                                         transforms.ToTensor(),
                                         normalize])

    img = Image.open(img_path)

    processed_img = data_transform(img)
    processed_img = processed_img.unsqueeze_(0)

    processed_img = Variable(processed_img)

    if use_cuda:
        processed_img = processed_img.cuda()

    prediction = VGG16(processed_img)
    if use_cuda:
        prediction = torch.argmax(prediction)
    else:
        prediction = prediction.data.cpu().numpy().argmax()

    return prediction # predicted class index
```

1.1.5 (IMPLEMENTATION) Write a Dog Detector

While looking at the dictionary, you will notice that the categories corresponding to dogs appear in an uninterrupted sequence and correspond to dictionary keys 151-268, inclusive, to include all categories from 'Chihuahua' to 'Mexican hairless'. Thus, in order to check to see if an image is predicted to contain a dog by the pre-trained VGG-16 model, we need only check if the pre-trained model predicts an index between 151 and 268 (inclusive).

Use these ideas to complete the dog_detector function below, which returns True if a dog is detected in an image (and False if not).

In [8]:
```
### returns "True" if a dog is detected in the image stored at img_path
def dog_detector(img_path):
    ## TODO: Complete the function.
    predicted_class_index = VGG16_predict(img_path)
    return predicted_class_index >= 151 and predicted_class_index <= 268
```

1.1.6 (IMPLEMENTATION) Assess the Dog Detector

Question 2: Use the code cell below to test the performance of your dog_detector function.
- What percentage of the images in human_files_short have a detected dog?
- What percentage of the images in dog_files_short have a detected dog?

Answer:

In [9]:
```
### TODO: Test the performance of the dog_detector function
### on the images in human_files_short and dog_files_short.


human_face_count = 0
dog_face_count = 0
for human_file in human_files_short:
    if dog_detector(human_file):
        human_face_count += 1
print('Percentage of human face detection', human_face_count / len(human_files_short))

for dog_file in dog_files_short:
    if dog_detector(dog_file):
        dog_face_count += 1
print('Percentage of dog face detection', dog_face_count / len(dog_files_short))
```

```
/opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/transforms/transf


Percentage of human face detection 0.0
Percentage of dog face detection 1.0
```

We suggest VGG-16 as a potential network to detect dog images in your algorithm, but you are free to explore other pre-trained networks (such as Inception-v3, ResNet-50, etc). Please use the code cell below to test other pre-trained PyTorch models. If you decide to pursue this *optional* task, report performance on human_files_short and dog_files_short.

```
In [10]: ### (Optional)
         ### TODO: Report the performance of another pre-trained network.
         ### Feel free to use as many code cells as needed.
```

Step 3: Create a CNN to Classify Dog Breeds (from Scratch)

Now that we have functions for detecting humans and dogs in images, we need a way to predict breed from images. In this step, you will create a CNN that classifies dog breeds. You must create your CNN *from scratch* (so, you can't use transfer learning *yet*!), and you must attain a test accuracy of at least 10%. In Step 4 of this notebook, you will have the opportunity to use transfer learning to create a CNN that attains greatly improved accuracy.

We mention that the task of assigning breed to dogs from images is considered exceptionally challenging. To see why, consider that *even a human* would have trouble distinguishing between a Brittany and a Welsh Springer Spaniel.

Brittany	Welsh Springer Spaniel

It is not difficult to find other dog breed pairs with minimal inter-class variation (for instance, Curly-Coated Retrievers and American Water Spaniels).

Curly-Coated Retriever	American Water Spaniel

Likewise, recall that labradors come in yellow, chocolate, and black. Your vision-based algorithm will have to conquer this high intra-class variation to determine how to classify all of these different shades as the same breed.

Yellow Labrador	Chocolate Labrador

We also mention that random chance presents an exceptionally low bar: setting aside the fact that the classes are slightly imabalanced, a random guess will provide a correct answer roughly 1 in 133 times, which corresponds to an accuracy of less than 1%.

Remember that the practice is far ahead of the theory in deep learning. Experiment with many different architectures, and trust your intuition. And, of course, have fun!

1.1.7 (IMPLEMENTATION) Specify Data Loaders for the Dog Dataset

Use the code cell below to write three separate data loaders for the training, validation, and test datasets of dog images (located at `dog_images/train`, `dog_images/valid`, and `dog_images/test`, respectively). You may find this documentation on custom datasets to be a useful resource. If you are interested in augmenting your training and/or validation data, check out the wide variety of transforms!

```
In [46]: import os
         from torchvision import datasets

         ### TODO: Write data loaders for training, validation, and test sets
         ## Specify appropriate transforms, and batch_sizes

         train_path = '/data/dog_images/train'
         valid_path = '/data/dog_images/valid'
         test_path = '/data/dog_images/test'

         loaders_scratch = {}

         normalize = transforms.Normalize(mean=[0.485, 0.456, 0.406],
                                          std=[0.229, 0.224, 0.225])

         data_transform_training = transforms.Compose([transforms.Scale(256),
                                              transforms.CenterCrop(224),
                                              transforms.RandomHorizontalFlip(),
                                              transforms.ToTensor(),
                                              normalize])

         data_transform_val_test = transforms.Compose([transforms.Scale(256),
                                              transforms.CenterCrop(224),
                                              transforms.ToTensor(),
                                              normalize])

         batch_size=32

         loaders_scratch['train'] = torch.utils.data.DataLoader(datasets.ImageFolder(train_path,
                                                        batch_size=batch_size, shuffle=T

         loaders_scratch['valid'] = torch.utils.data.DataLoader(datasets.ImageFolder(valid_path,

         loaders_scratch['test'] = torch.utils.data.DataLoader(datasets.ImageFolder(test_path, d

/opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/transforms/transf
```

Question 3: Describe your chosen procedure for preprocessing the data. - How does your code resize the images (by cropping, stretching, etc)? What size did you pick for the input tensor, and why? - Did you decide to augment the dataset? If so, how (through translations, flips, rotations, etc)? If not, why not?

Answer: centercrop to 224 *224 image since vgg accepts the image of this size used randomRotation to augment the datasetm so as to increase the model's ability to learn the invariants representation of the image

1.1.8 (IMPLEMENTATION) Model Architecture

Create a CNN to classify dog breed. Use the template in the code cell below.

Question 4: Outline the steps you took to get to your final CNN architecture and your reasoning at each step.

```
In [48]: import torch.nn as nn
         import torch.nn.functional as F

         # define the CNN architecture
         class Net(nn.Module):
             ### TODO: choose an architecture, and complete the class
             def __init__(self):
                 super(Net, self).__init__()
                 ## Define layers of a CNN

                 self.conv1 = nn.Conv2d(3, 64, 7, 2, padding=3)
                 self.conv2 = nn.Conv2d(64, 64, 3, padding=1)
                 self.conv3 = nn.Conv2d(64, 128, 3, 2, padding=1)
                 self.conv4 = nn.Conv2d(128, 128, 3, padding=1)
                 self.conv5 = nn.Conv2d(128, 256, 3, 2, padding=1)
                 self.conv6 = nn.Conv2d(256, 256, 3, padding=1)

                 self.conv7 = nn.Conv2d(256, 512, 3, padding=1)


                 self.pool_1 = nn.MaxPool2d(3, stride = 2, padding =1)
                 self.pool_2 = nn.MaxPool2d(2, 2)

                 self.fc1 = nn.Linear(512 , 133)

                 self.bn2_1 = nn.BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_run

                 self.bn2_2 = nn.BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_ru

                 self.bn2_3 = nn.BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_ru

                 self.bn2_4 = nn.BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_ru



                 self.dropout = nn.Dropout(0.25)

                 self.avg = nn.AvgPool2d(7)

             def forward(self, x):
                 ## Define forward behavior
                 x = self.conv1(x)

                 x = self.pool_1(F.relu(self.bn2_1(self.conv2(x))))
```

```
            x = self.conv3(x)

            x = self.pool_1(self.bn2_2(F.relu(self.conv4(x))))


            x = self.conv5(x)

            x = self.bn2_3(F.relu(self.conv6(x)))


#             x = self.pool_1(F.relu(self.conv3(x)))

#             x = F.relu(self.bn2_2(self.conv4(x)))

#             x = self.pool_1(F.relu(self.conv5(x)))

#             x = self.pool_1(self.bn2_3(F.relu(self.conv6(x))))

            x = self.avg(F.relu(self.bn2_4(self.conv7(x))))

            #x = self.dropout(x)
            x = x.view(-1, 512 * 1 * 1)



            #x = F.relu(self.fc1(x))

            #x = self.dropout(x)

            x = self.fc1(x)

            return x

#-#-# You so NOT have to modify the code below this line. #-#-#

# instantiate the CNN
model_scratch = Net()


# move tensors to GPU if CUDA is available

use_cuda = torch.cuda.is_available()

if use_cuda:
    model_scratch.cuda()
```

In [24]:
```
model_scratch = models.resnet18()
```

```python
print(model_scratch)
n_inputs = model_scratch.fc.in_features

last_layer = nn.Linear(n_inputs, 133)
model_scratch.fc = last_layer


if use_cuda:
    model_scratch = model_scratch.cuda()
```

```
ResNet(
  (conv1): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace)
  (maxpool): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1, ceil_mode=False)
  (layer1): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
    (1): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (layer2): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(64, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (downsample): Sequential(
        (0): Conv2d(64, 128, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
```

```
      (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (layer3): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (downsample): Sequential(
        (0): Conv2d(128, 256, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (layer4): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(256, 512, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (downsample): Sequential(
        (0): Conv2d(256, 512, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace)
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (avgpool): AvgPool2d(kernel_size=7, stride=1, padding=0)
  (fc): Linear(in_features=512, out_features=1000, bias=True)
)
```

```
In [20]: model_scratch = models.vgg16()


         n_inputs = model_scratch.classifier[6].in_features

         print(n_inputs)

         last_layer = nn.Linear(n_inputs, 133)
         model_scratch.classifier[6] = last_layer


         if use_cuda:
             model_scratch = model_scratch.cuda()
```

```
4096
```

Answer:

1.1.9 (IMPLEMENTATION) Specify Loss Function and Optimizer

Use the next code cell to specify a loss function and optimizer. Save the chosen loss function as `criterion_scratch`, and the optimizer as `optimizer_scratch` below.

```
In [30]: import torch.optim as optim

         ### TODO: select loss function
         criterion_scratch = nn.CrossEntropyLoss()

         ### TODO: select optimizer
         optimizer_scratch = optim.Adam(model_scratch.parameters(), lr=0.01)
```

```
In [31]: from matplotlib import pyplot as plt
```

1.1.10 (IMPLEMENTATION) Train and Validate the Model

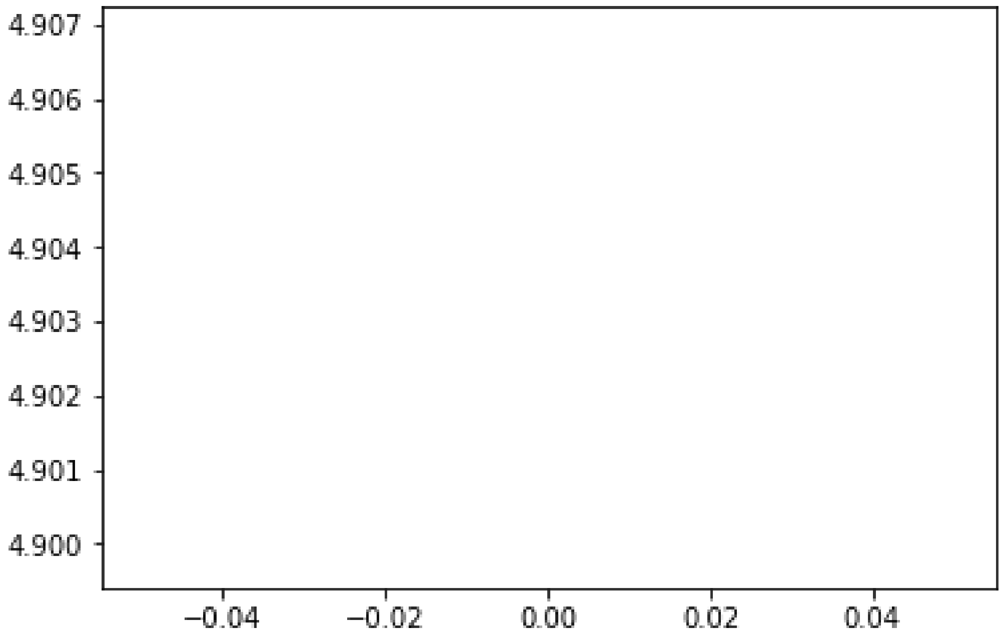
Train and validate your model in the code cell below. Save the final model parameters at filepath `'model_scratch.pt'`.

```
In [49]: from PIL import ImageFile
         ImageFile.LOAD_TRUNCATED_IMAGES = True

         train_losses = []
         valid_losses = []

         def train(n_epochs, loaders, model, optimizer, criterion, use_cuda, save_path):
             """returns trained model"""
             # initialize tracker for minimum validation xoss
             valid_loss_min = np.Inf
```

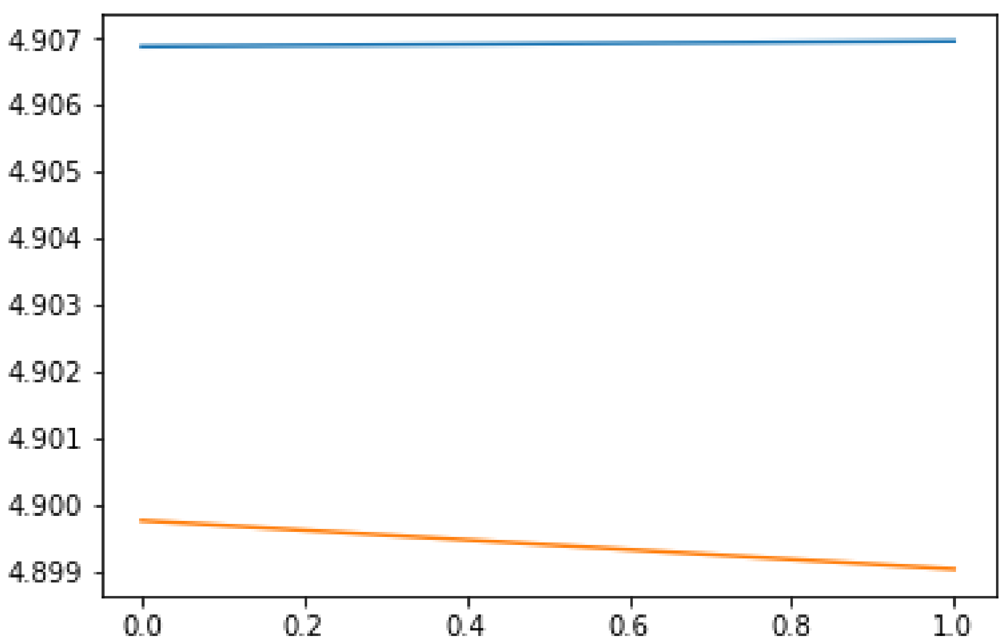
```
for epoch in range(1, n_epochs+1):
    # initialize variables to monitor training and validation loss
    train_loss = 0.0
    valid_loss = 0.0

    ###################
    # train the model #
    ###################
    model.train()
    for batch_idx, (data, target) in enumerate(loaders['train']):
        # move to GPU
        if use_cuda:
            data, target = data.cuda(), target.cuda()
        ## find the loss and update the model parameters accordingly
        ## record the average training loss, using something like
        ## train_loss = train_loss + ((1 / (batch_idx + 1)) * (loss.data - train_lo
        optimizer.zero_grad()
        output = model(data)
        loss = criterion(output, target)
        loss.backward()
        optimizer.step()
        train_loss = train_loss + loss.data

    train_loss = train_loss / (batch_idx + 1)
    train_losses.append(train_loss)

    ######################
    # validate the model #
    ######################
    model.eval()
    for batch_idx, (data, target) in enumerate(loaders['valid']):
        # move to GPU
        if use_cuda:
            data, target = data.cuda(), target.cuda()
        ## update the average validation loss
        output = model(data)
        loss = criterion(output, target)

        valid_loss = valid_loss + loss.data
    valid_loss = valid_loss / (batch_idx + 1 )
    valid_losses.append(valid_loss)
    # print training/validation statistics
    print('Epoch: {} \tTraining Loss: {:.6f} \tValidation Loss: {:.6f}'.format(
        epoch,
        train_loss,
        valid_loss
        ))
```

```
            # TODO: save the model if validation loss has decreased

            if valid_loss <= valid_loss_min:
                print('Validation loss decreased ({:.6f} --> {:.6f}).  Saving model ...'.fo
                valid_loss_min,
                valid_loss))
                torch.save(model.state_dict(), save_path)
                valid_loss_min = valid_loss

            plt.plot(train_losses)
            plt.plot(valid_losses)
            plt.show()

        # return trained model
        return model


    #train the model

    save_path = 'model_scratch.pt'
    model_scratch = train(100, loaders_scratch, model_scratch, optimizer_scratch,
                          criterion_scratch, use_cuda, save_path)

    #load the model that got the best validation accuracy
    model_scratch.load_state_dict(torch.load('model_scratch.pt'))
```

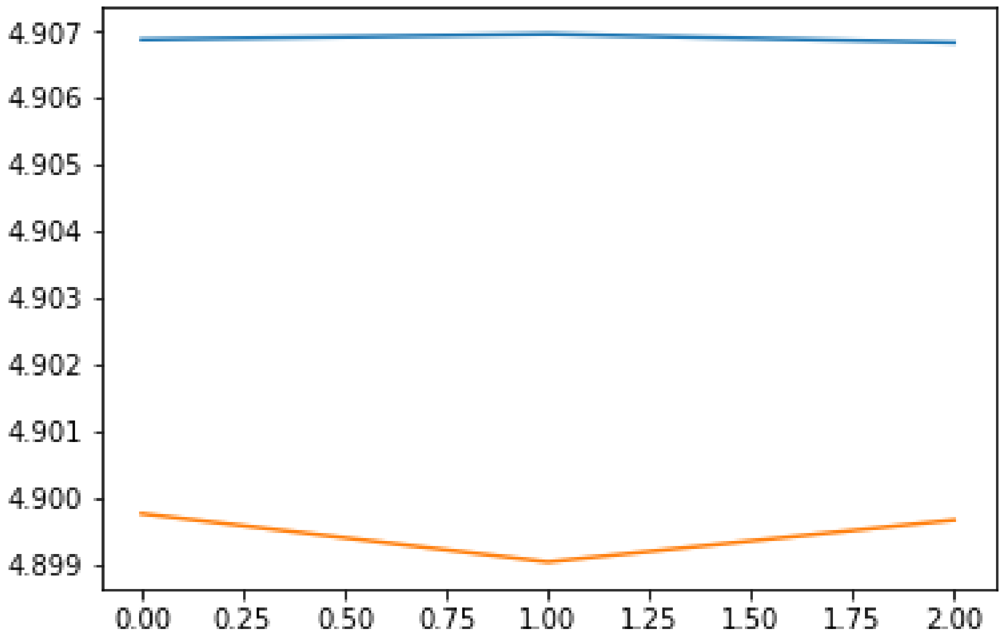
```
Epoch: 1        Training Loss: 4.906876         Validation Loss: 4.899763
Validation loss decreased (inf --> 4.899763).  Saving model ...
```



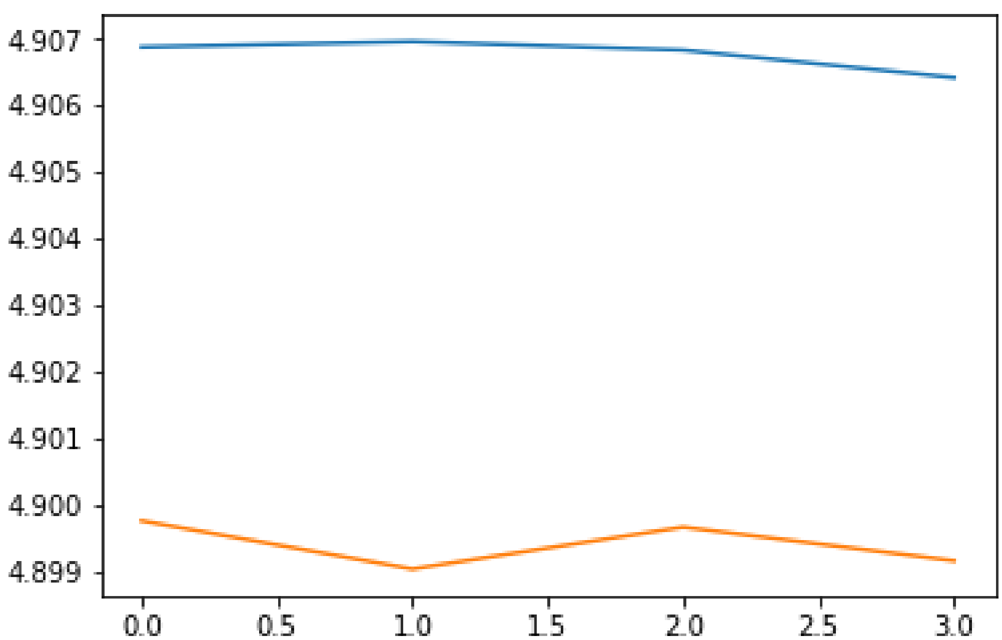

```
Epoch: 2        Training Loss: 4.906958         Validation Loss: 4.899037
Validation loss decreased (4.899763 --> 4.899037).  Saving model ...
```



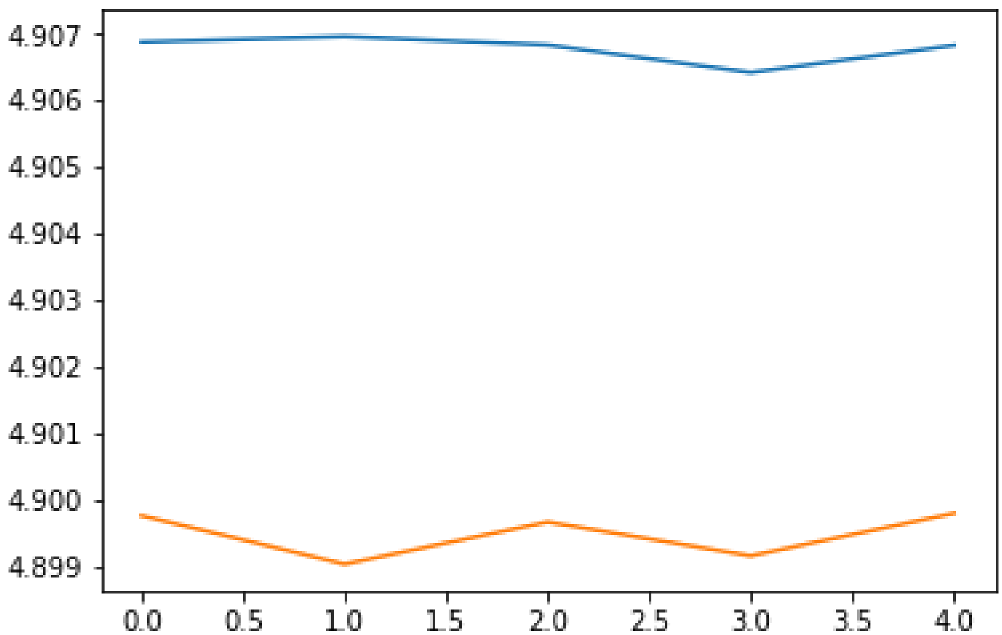
```
Epoch: 3 	Training Loss: 4.906831 	Validation Loss: 4.899667
```



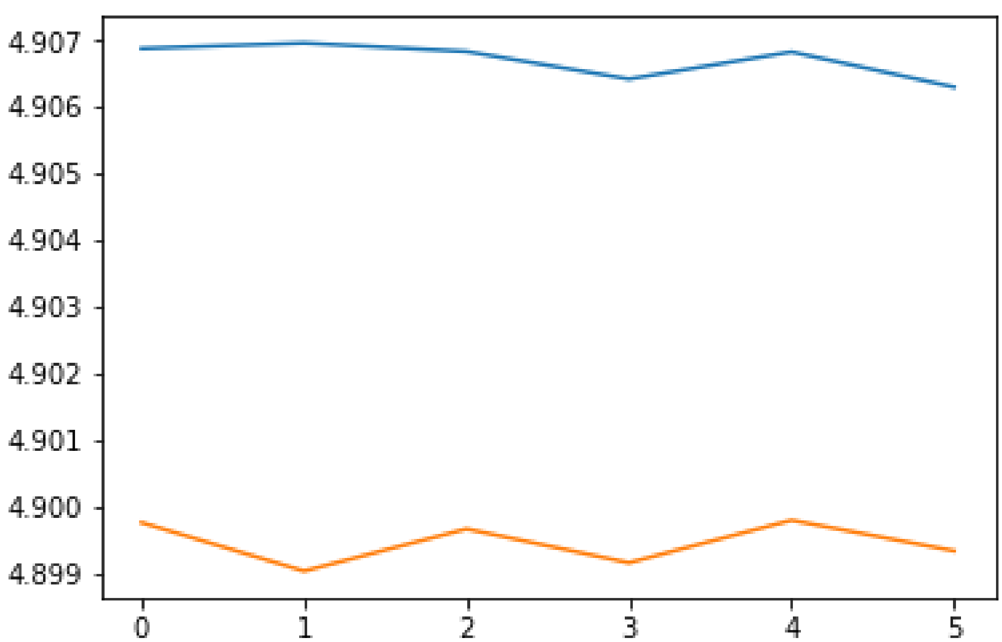

```
Epoch: 4 	Training Loss: 4.906417 	Validation Loss: 4.899162
```



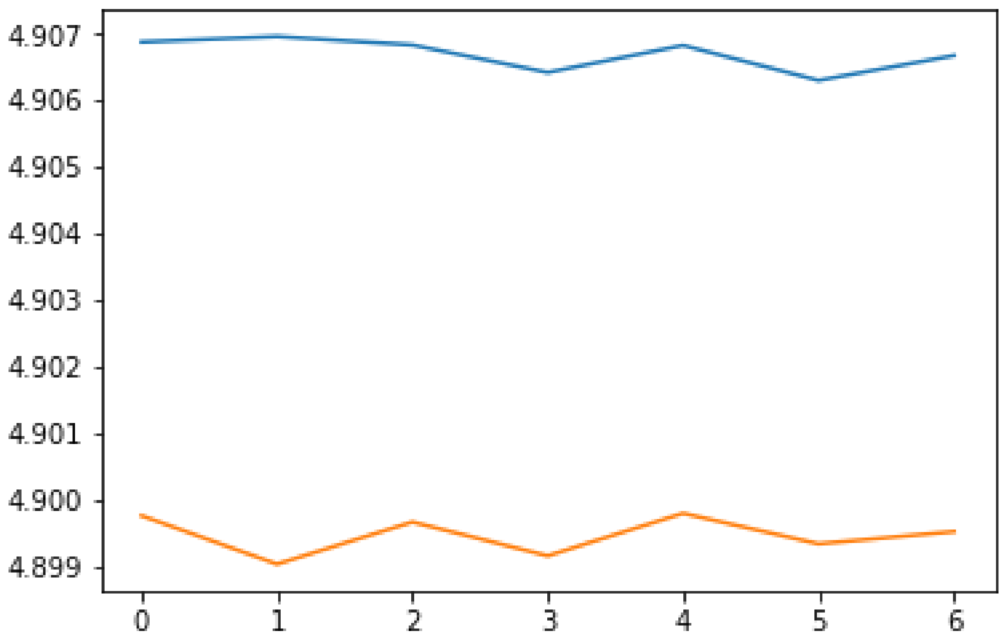
```
Epoch: 5          Training Loss: 4.906825          Validation Loss: 4.899798
```



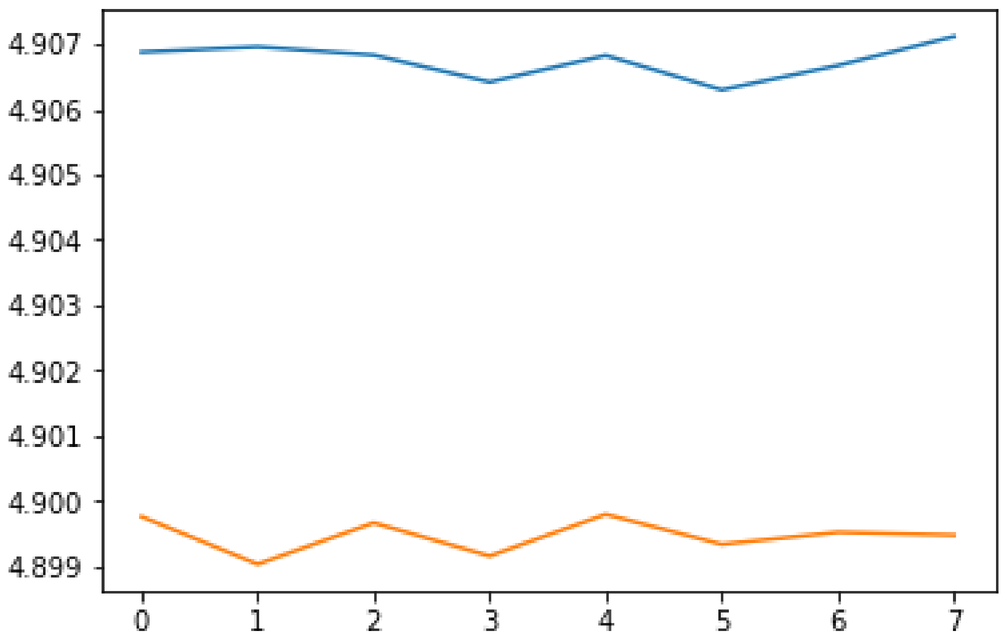

```
Epoch: 6          Training Loss: 4.906298          Validation Loss: 4.899340
```



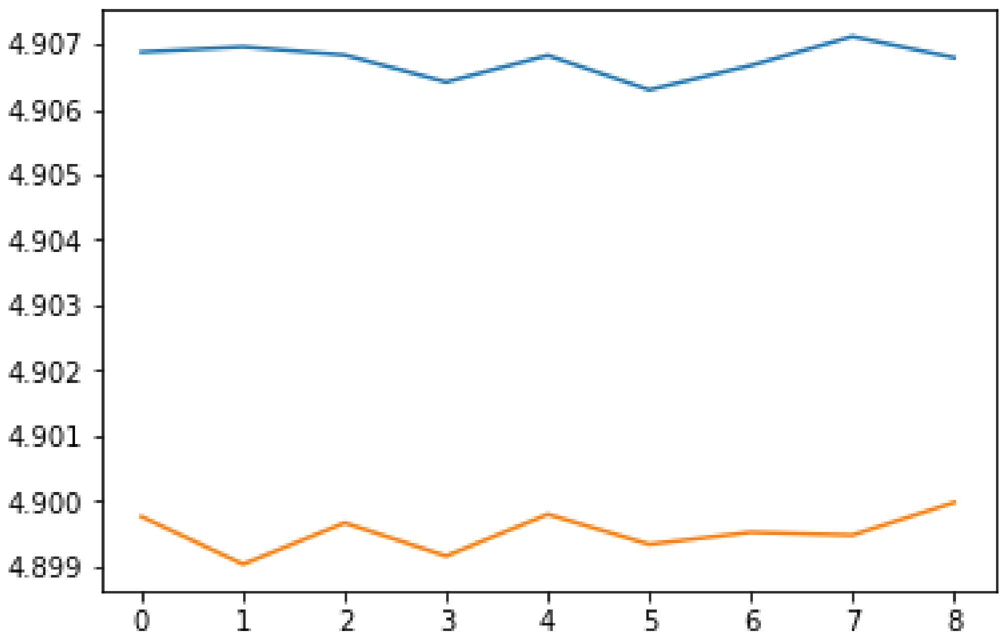
```
Epoch: 7         Training Loss: 4.906671         Validation Loss: 4.899523
```



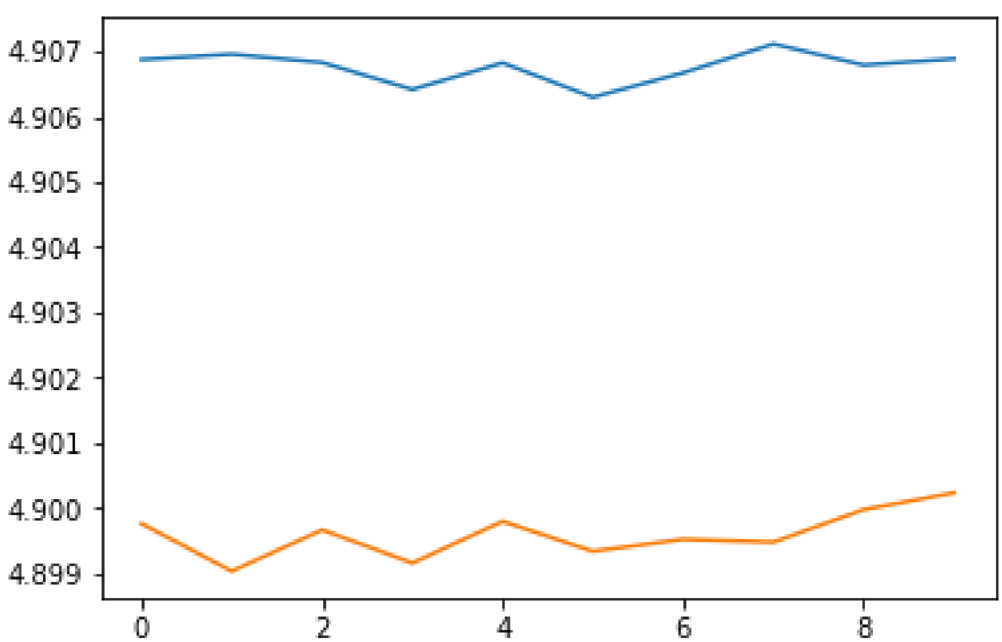

```
Epoch: 8         Training Loss: 4.907115         Validation Loss: 4.899480
```



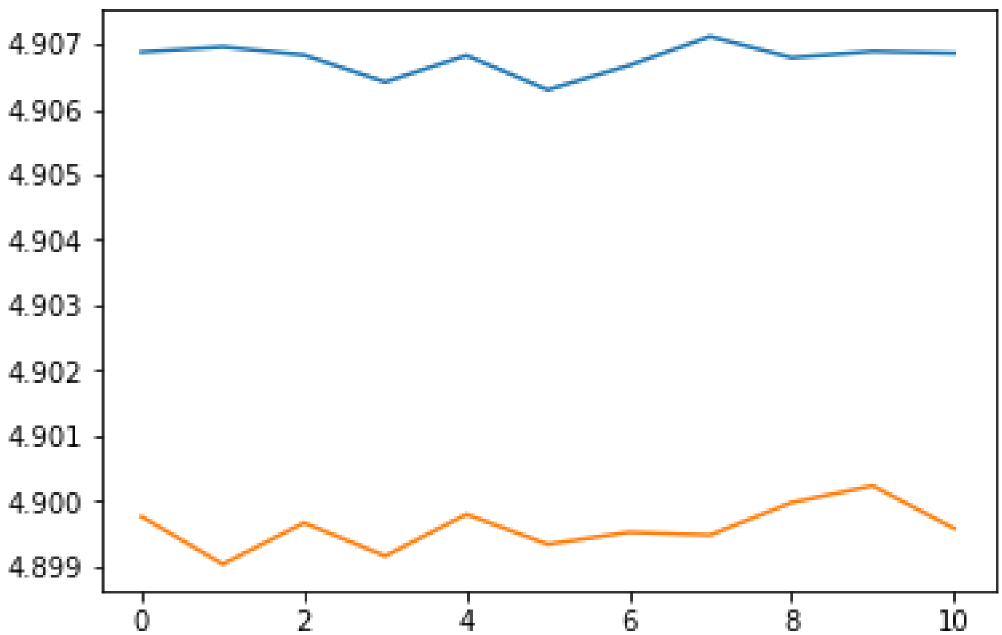
```
Epoch: 9          Training Loss: 4.906792          Validation Loss: 4.899979
```



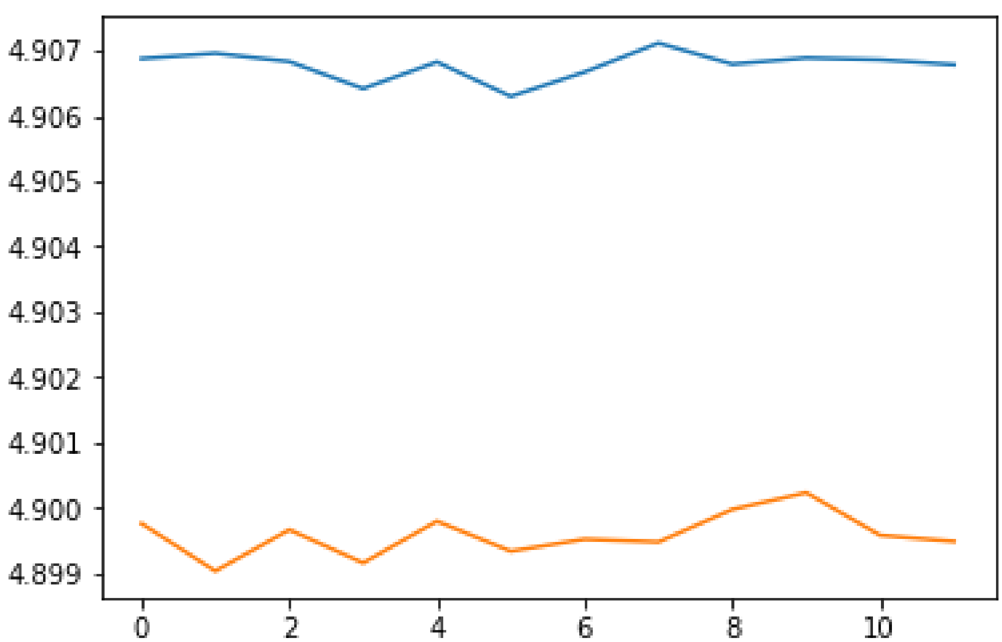

```
Epoch: 10         Training Loss: 4.906887          Validation Loss: 4.900236
```



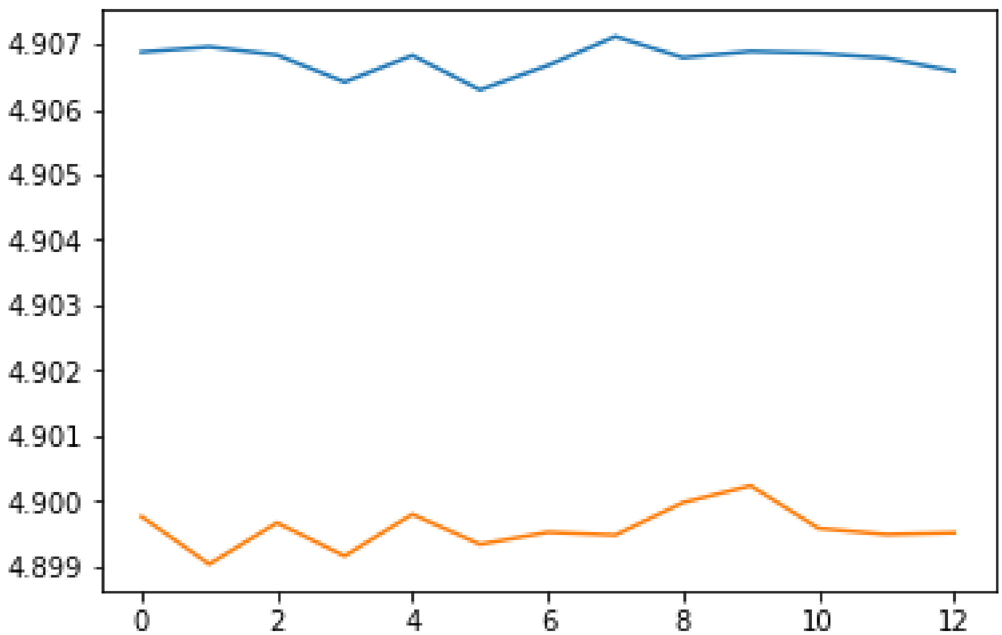
Epoch: 11 Training Loss: 4.906859 Validation Loss: 4.899583



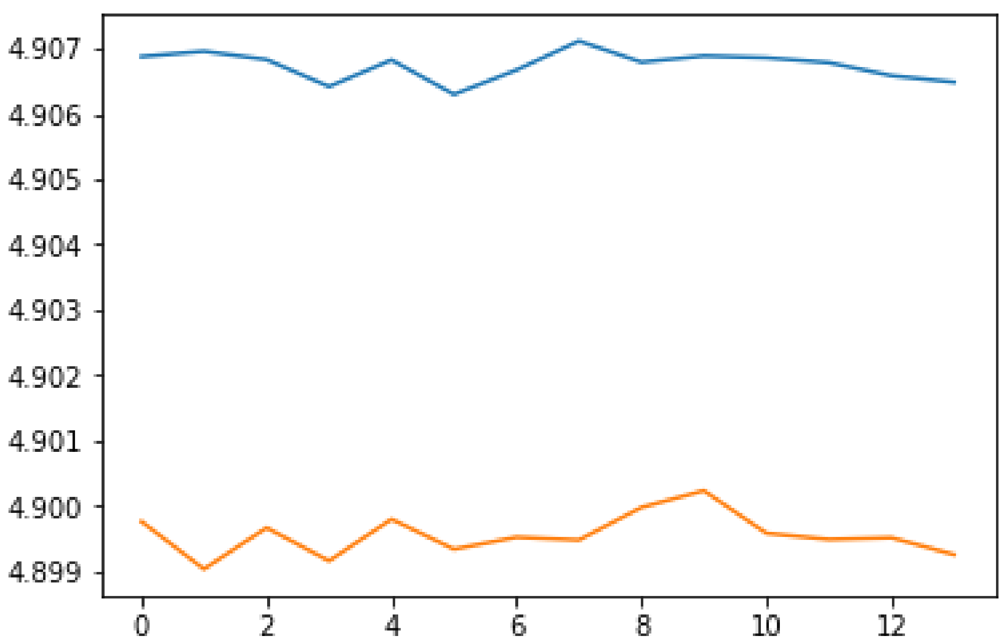
Epoch: 12 Training Loss: 4.906784 Validation Loss: 4.899489



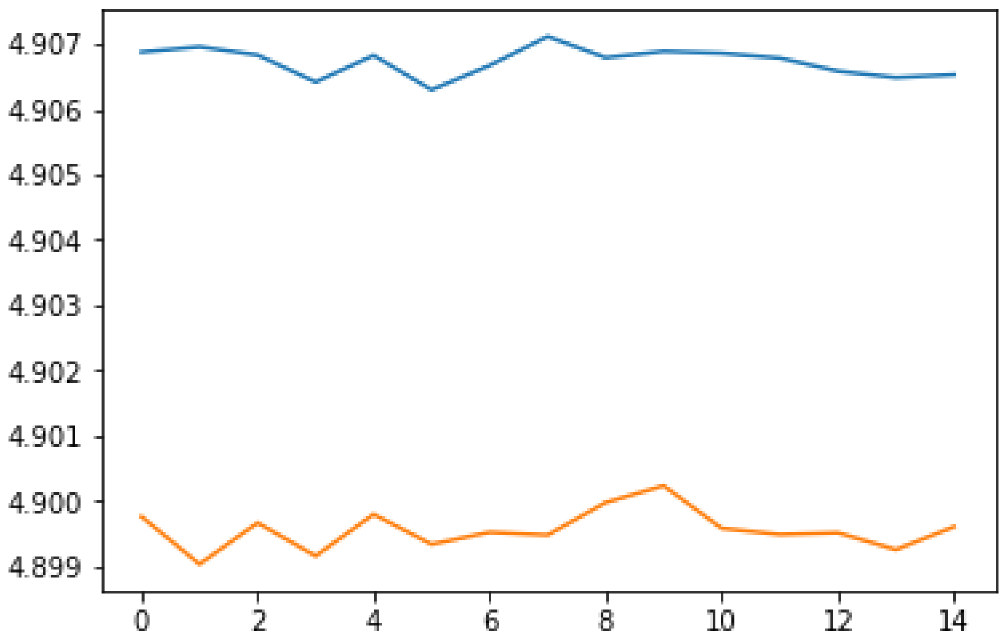
Epoch: 13 Training Loss: 4.906584 Validation Loss: 4.899515



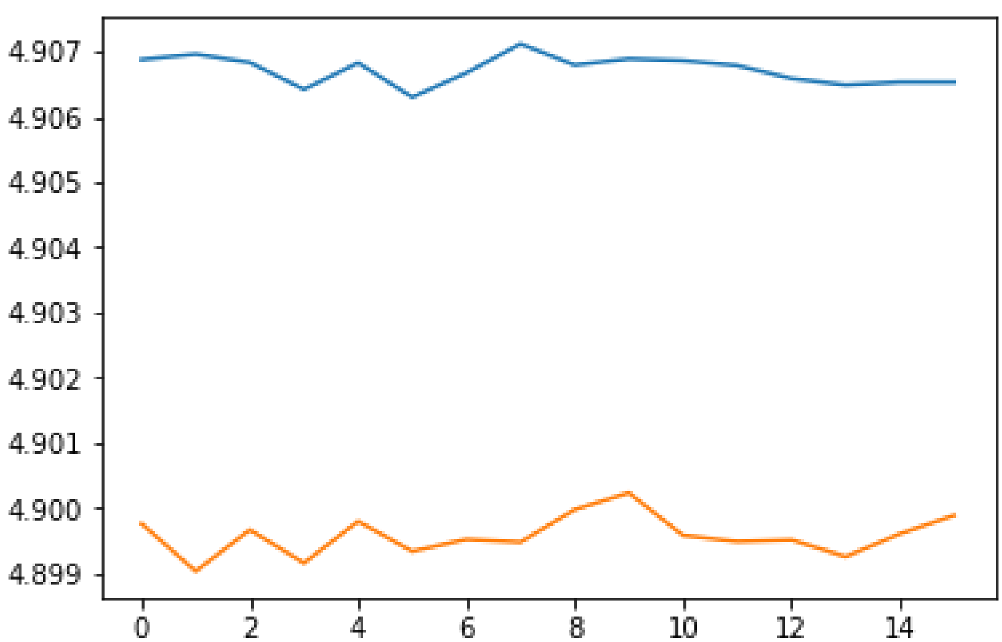
Epoch: 14 Training Loss: 4.906486 Validation Loss: 4.899259



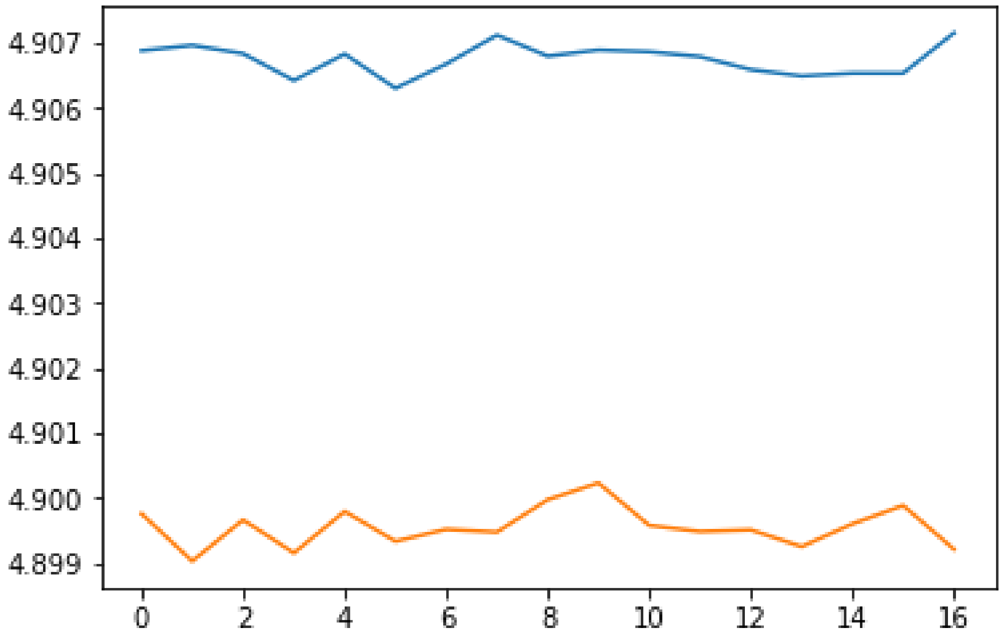
```
Epoch: 15         Training Loss: 4.906529         Validation Loss: 4.899606
```



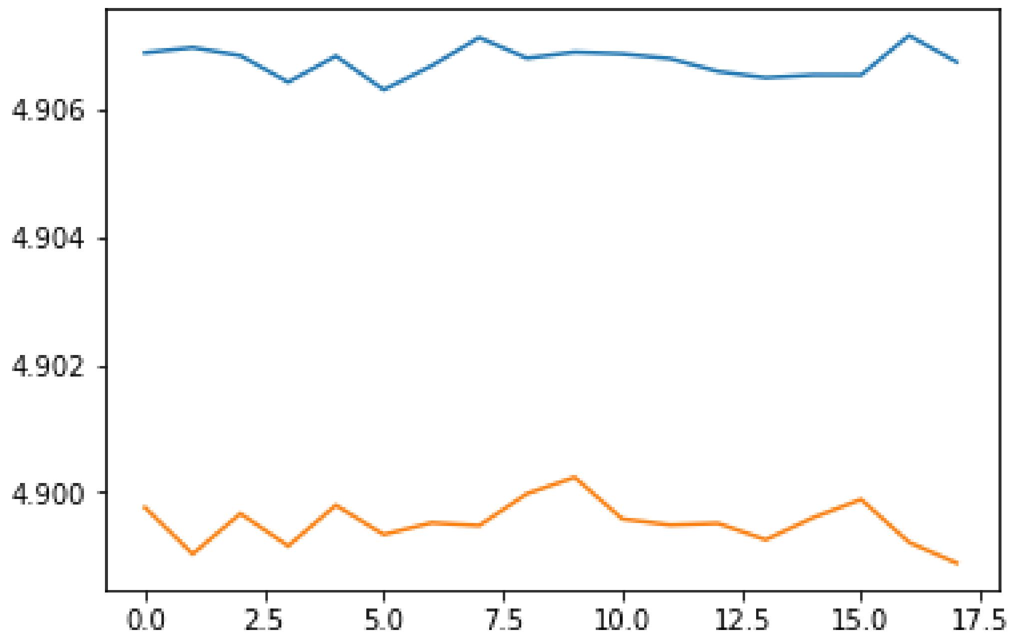

```
Epoch: 16         Training Loss: 4.906528         Validation Loss: 4.899889
```



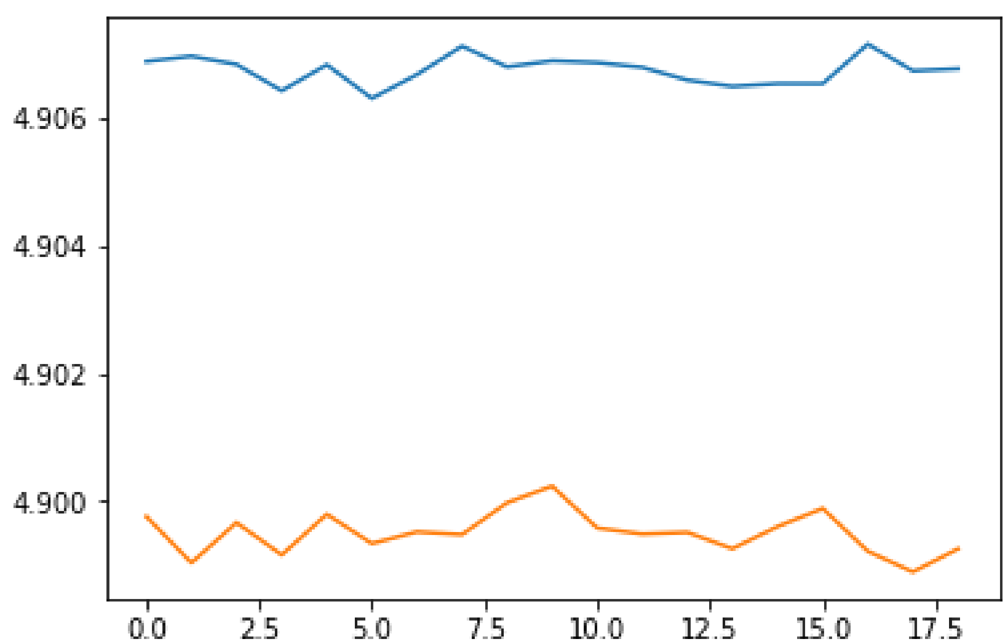
```
Epoch: 17          Training Loss: 4.907147          Validation Loss: 4.899220
```




```
Epoch: 18          Training Loss: 4.906731          Validation Loss: 4.898895
Validation loss decreased (4.899037 --> 4.898895).  Saving model ...
```




```
Epoch: 19          Training Loss: 4.906762          Validation Loss: 4.899256
```



```
---------------------------------------------------------------------------

KeyboardInterrupt                         Traceback (most recent call last)

<ipython-input-49-329e8eebdce3> in <module>()
     80 save_path = 'model_scratch.pt'
     81 model_scratch = train(100, loaders_scratch, model_scratch, optimizer_scratch,
---> 82                       criterion_scratch, use_cuda, save_path)
     83 
     84 #load the model that got the best validation accuracy


<ipython-input-49-329e8eebdce3> in train(n_epochs, loaders, model, optimizer, criterion,
     41         ######################
     42         model.eval()
---> 43         for batch_idx, (data, target) in enumerate(loaders['valid']):
     44             # move to GPU
     45             if use_cuda:


/opt/conda/lib/python3.6/site-packages/torch/utils/data/dataloader.py in __next__(self)
    262         if self.num_workers == 0:  # same-process loading
    263             indices = next(self.sample_iter)  # may raise StopIteration
--> 264             batch = self.collate_fn([self.dataset[i] for i in indices])
    265             if self.pin_memory:
    266                 batch = pin_memory_batch(batch)


/opt/conda/lib/python3.6/site-packages/torch/utils/data/dataloader.py in <listcomp>(.0)
    262         if self.num_workers == 0:  # same-process loading
    263             indices = next(self.sample_iter)  # may raise StopIteration
--> 264             batch = self.collate_fn([self.dataset[i] for i in indices])
    265             if self.pin_memory:
    266                 batch = pin_memory_batch(batch)


/opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/datasets/
     99         """
    100         path, target = self.samples[index]
--> 101         sample = self.loader(path)
    102         if self.transform is not None:
    103             sample = self.transform(sample)


/opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/datasets/
    145         return accimage_loader(path)
    146     else:
```

```
--> 147         return pil_loader(path)
    148
    149


    /opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/datasets/
    128     with open(path, 'rb') as f:
    129         img = Image.open(f)
--> 130         return img.convert('RGB')
    131
    132


    /opt/conda/lib/python3.6/site-packages/PIL/Image.py in convert(self, mode, matrix, dithe
    890         """
    891
--> 892         self.load()
    893
    894         if not mode and self.mode == "P":


    /opt/conda/lib/python3.6/site-packages/PIL/ImageFile.py in load(self)
    192                     self.map = None
    193
--> 194         self.load_prepare()
    195         err_code = -3  # initialize to unknown error
    196         if not self.map:


    /opt/conda/lib/python3.6/site-packages/PIL/ImageFile.py in load_prepare(self)
    260         if not self.im or\
    261            self.im.mode != self.mode or self.im.size != self.size:
--> 262             self.im = Image.core.new(self.mode, self.size)
    263         # create palette (optional)
    264         if self.mode == "P":


    KeyboardInterrupt:
```

In [17]:
```
torch.__version__
```

Out[17]: '0.4.0'

In [50]:
```
#load the model that got the best validation accuracy
model_scratch.load_state_dict(torch.load('model_scratch.pt'))
```

1.1.11 (IMPLEMENTATION) Test the Model

Try out your model on the test dataset of dog images. Use the code cell below to calculate and print the test loss and accuracy. Ensure that your test accuracy is greater than 10%.

```
In [51]: def test(loaders, model, criterion, use_cuda):

             # monitor test loss and accuracy
             test_loss = 0.
             correct = 0.
             total = 0.

             model.eval()
             for batch_idx, (data, target) in enumerate(loaders['test']):
                 # move to GPU
                 if use_cuda:
                     data, target = data.cuda(), target.cuda()
                 # forward pass: compute predicted outputs by passing inputs to the model
                 output = model(data)
                 # calculate the loss
                 loss = criterion(output, target)
                 # update average test loss
                 test_loss = test_loss + ((1 / (batch_idx + 1)) * (loss.data - test_loss))
                 # convert output probabilities to predicted class
                 pred = output.data.max(1, keepdim=True)[1]
                 # compare predictions to true label
                 correct += np.sum(np.squeeze(pred.eq(target.data.view_as(pred))).cpu().numpy())
                 total += data.size(0)

             print('Test Loss: {:.6f}\n'.format(test_loss))

             print('\nTest Accuracy: %2d%% (%2d/%2d)' % (
                 100. * correct / total, correct, total))

         # call test function
         test(loaders_scratch, model_scratch, criterion_scratch, use_cuda)
```

```
Test Loss: 4.894906


Test Accuracy:  0% ( 7/836)
```

Step 4: Create a CNN to Classify Dog Breeds (using Transfer Learning)

You will now use transfer learning to create a CNN that can identify dog breed from images. Your CNN must attain at least 60% accuracy on the test set.

1.1.12 (IMPLEMENTATION) Specify Data Loaders for the Dog Dataset

Use the code cell below to write three separate data loaders for the training, validation, and test datasets of dog images (located at `dogImages/train`, `dogImages/valid`, and `dogImages/test`, respectively).

If you like, **you are welcome to use the same data loaders from the previous step**, when you created a CNN from scratch.

```
In [54]: ## TODO: Specify data loaders

         import os
         from torchvision import datasets

         ### TODO: Write data loaders for training, validation, and test sets
         ## Specify appropriate transforms, and batch_sizes

         train_path = '/data/dog_images/train'
         valid_path = '/data/dog_images/valid'
         test_path = '/data/dog_images/test'

         loaders_transfer = {}

         normalize = transforms.Normalize(mean=[0.485, 0.456, 0.406],
                                          std=[0.229, 0.224, 0.225])

         data_transform = transforms.Compose([transforms.Scale(256),
                                              transforms.CenterCrop(224),
                                              transforms.RandomRotation(30),
                                              transforms.ToTensor(),
                                              normalize])

         loaders_transfer['train'] = torch.utils.data.DataLoader(datasets.ImageFolder(train_path

         loaders_transfer['valid'] = torch.utils.data.DataLoader(datasets.ImageFolder(valid_path

         loaders_transfer['test'] = torch.utils.data.DataLoader(datasets.ImageFolder(test_path,
```

```
/opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/transforms/transf
  "please use transforms.Resize instead.")
```

1.1.13 (IMPLEMENTATION) Model Architecture

Use transfer learning to create a CNN to classify dog breed. Use the code cell below, and save your initialized model as the variable `model_transfer`.

```
In [58]: import torchvision.models as models
         import torch.nn as nn
```

```
#models.resnet18, models.resnet51


## TODO: Specify model architecture

model_transfer = models.vgg16(pretrained=True)

for param in model_transfer.features.parameters():
    param.requires_grad = False

n_inputs = model_transfer.classifier[6].in_features

print(n_inputs)

last_layer = nn.Linear(n_inputs, 133)
model_transfer.classifier[6] = last_layer


if use_cuda:
    model_transfer = model_transfer.cuda()
```

```
4096
```

```
In [ ]: (conv -> conv -> relu -> batch_norm -> pool)+

        or


        (conv -> conv -> relu -> pool) +
```

Question 5: Outline the steps you took to get to your final CNN architecture and your reasoning at each step. Describe why you think the architecture is suitable for the current problem.
Answer:

1.1.14 (IMPLEMENTATION) Specify Loss Function and Optimizer

Use the next code cell to specify a loss function and optimizer. Save the chosen loss function as `criterion_transfer`, and the optimizer as `optimizer_transfer` below.

```
In [59]: criterion_transfer = nn.CrossEntropyLoss()
         optimizer_transfer = optim.SGD(model_transfer.classifier.parameters(), lr=0.001)
```

1.1.15 (IMPLEMENTATION) Train and Validate the Model

Train and validate your model in the code cell below. Save the final model parameters at filepath `'model_transfer.pt'`.

In [62]:
```
# train the model
model_transfer = train(100, loaders_transfer, model_transfer, optimizer_transfer, crite

# load the model that got the best validation accuracy (uncomment the line below)
model_transfer.load_state_dict(torch.load('model_transfer.pt'))
```

```
Epoch: 1         Training Loss: 4.746706        Validation Loss: 4.276037
Validation loss decreased (inf --> 4.276037).  Saving model ...
Epoch: 2         Training Loss: 4.385982        Validation Loss: 3.948139
Validation loss decreased (4.276037 --> 3.948139).  Saving model ...
Epoch: 3         Training Loss: 4.066204        Validation Loss: 3.569814
Validation loss decreased (3.948139 --> 3.569814).  Saving model ...
Epoch: 4         Training Loss: 3.727339        Validation Loss: 3.227735
Validation loss decreased (3.569814 --> 3.227735).  Saving model ...
Epoch: 5         Training Loss: 3.372212        Validation Loss: 2.817487
Validation loss decreased (3.227735 --> 2.817487).  Saving model ...
Epoch: 6         Training Loss: 3.037870        Validation Loss: 2.483777
Validation loss decreased (2.817487 --> 2.483777).  Saving model ...
Epoch: 7         Training Loss: 2.713423        Validation Loss: 2.166200
Validation loss decreased (2.483777 --> 2.166200).  Saving model ...
Epoch: 8         Training Loss: 2.482719        Validation Loss: 1.913581
Validation loss decreased (2.166200 --> 1.913581).  Saving model ...
Epoch: 9         Training Loss: 2.264276        Validation Loss: 1.693540
Validation loss decreased (1.913581 --> 1.693540).  Saving model ...
Epoch: 10         Training Loss: 2.096090        Validation Loss: 1.543260
Validation loss decreased (1.693540 --> 1.543260).  Saving model ...
Epoch: 11         Training Loss: 1.939055        Validation Loss: 1.403392
Validation loss decreased (1.543260 --> 1.403392).  Saving model ...
Epoch: 12         Training Loss: 1.829363        Validation Loss: 1.307422
Validation loss decreased (1.403392 --> 1.307422).  Saving model ...
Epoch: 13         Training Loss: 1.704023        Validation Loss: 1.249190
Validation loss decreased (1.307422 --> 1.249190).  Saving model ...
Epoch: 14         Training Loss: 1.642562        Validation Loss: 1.205281
Validation loss decreased (1.249190 --> 1.205281).  Saving model ...
```

```
---------------------------------------------------------------------------

KeyboardInterrupt                         Traceback (most recent call last)

<ipython-input-62-fb32d09aecd6> in <module>()
      1 # train the model
----> 2 model_transfer = train(100, loaders_transfer, model_transfer, optimizer_transfer, cr
      3
      4 # load the model that got the best validation accuracy (uncomment the line below)
      5 model_transfer.load_state_dict(torch.load('model_transfer.pt'))
```

```
<ipython-input-50-6c753e00cfc6> in train(n_epochs, loaders, model, optimizer, criterion,
     17         ###################
     18         model.train()
---> 19         for batch_idx, (data, target) in enumerate(loaders['train']):
     20             # move to GPU
     21             if use_cuda:


/opt/conda/lib/python3.6/site-packages/torch/utils/data/dataloader.py in __next__(self)
    262         if self.num_workers == 0:  # same-process loading
    263             indices = next(self.sample_iter)  # may raise StopIteration
--> 264             batch = self.collate_fn([self.dataset[i] for i in indices])
    265             if self.pin_memory:
    266                 batch = pin_memory_batch(batch)


/opt/conda/lib/python3.6/site-packages/torch/utils/data/dataloader.py in <listcomp>(.0)
    262         if self.num_workers == 0:  # same-process loading
    263             indices = next(self.sample_iter)  # may raise StopIteration
--> 264             batch = self.collate_fn([self.dataset[i] for i in indices])
    265             if self.pin_memory:
    266                 batch = pin_memory_batch(batch)


/opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/datasets/
     99         """
    100         path, target = self.samples[index]
--> 101         sample = self.loader(path)
    102         if self.transform is not None:
    103             sample = self.transform(sample)


/opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/datasets/
    145         return accimage_loader(path)
    146     else:
--> 147         return pil_loader(path)
    148
    149


/opt/conda/lib/python3.6/site-packages/torchvision-0.2.1-py3.6.egg/torchvision/datasets/
    128     with open(path, 'rb') as f:
    129         img = Image.open(f)
--> 130         return img.convert('RGB')
    131
    132
```

```
/opt/conda/lib/python3.6/site-packages/PIL/Image.py in convert(self, mode, matrix, dithe
890         """
891
--> 892         self.load()
893
894         if not mode and self.mode == "P":


/opt/conda/lib/python3.6/site-packages/PIL/ImageFile.py in load(self)
233
234                             b = b + s
--> 235                             n, err_code = decoder.decode(b)
236                             if n < 0:
237                                 break


KeyboardInterrupt:
```

In [63]:
```python
# load the model that got the best validation accuracy (uncomment the line below)
model_transfer.load_state_dict(torch.load('model_transfer.pt'))
```

1.1.16 (IMPLEMENTATION) Test the Model

Try out your model on the test dataset of dog images. Use the code cell below to calculate and print the test loss and accuracy. Ensure that your test accuracy is greater than 60%.

In [64]:
```python
test(loaders_transfer, model_transfer, criterion_transfer, use_cuda)
```

```
Test Loss: 1.202827


Test Accuracy: 70% (589/836)
```

1.1.17 (IMPLEMENTATION) Predict Dog Breed with the Model

Write a function that takes an image path as input and returns the dog breed (`Affenpinscher`, `Afghan hound`, etc) that is predicted by your model.

In [ ]:
```python
### TODO: Write a function that takes a path to an image as input
### and returns the dog breed that is predicted by the model.

# list of class names by index, i.e. a name can be accessed like class_names[0]
class_names = [item[4:].replace("_", " ") for item in data_transfer['train'].classes]

def predict_breed_transfer(img_path):
    # load the image and return the predicted breed
```

```
    img = Image.open(img_path)

    normalize = transforms.Normalize(mean=[0.485, 0.456, 0.406],
                                     std=[0.229, 0.224, 0.225])

    data_transform = transforms.Compose([transforms.Scale(256),
                                         transforms.CenterCrop(224),
                                         transforms.ToTensor(),
                                         normalize])

    img = Image.open(img_path)

    processed_img = data_transform(img)
    processed_img = processed_img.unsqueeze_(0)

    processed_img = Variable(processed_img)

    if use_cuda:
        processed_img = processed_img.cuda()

    output = model_transfer(processed_img)

    prediction = output.data.max(1, keepdim=True)[1]


    return prediction
```

Step 5: Write your Algorithm

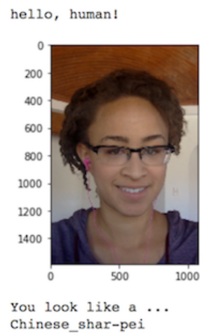
Write an algorithm that accepts a file path to an image and first determines whether the image contains a human, dog, or neither. Then, - if a **dog** is detected in the image, return the predicted breed. - if a **human** is detected in the image, return the resembling dog breed. - if **neither** is detected in the image, provide output that indicates an error.

You are welcome to write your own functions for detecting humans and dogs in images, but feel free to use the `face_detector` and `human_detector` functions developed above. You are **required** to use your CNN from Step 4 to predict dog breed.

Some sample output for our algorithm is provided below, but feel free to design your own user experience!

1.1.18 (IMPLEMENTATION) Write your Algorithm

```
In [ ]: ### TODO: Write your algorithm.
        ### Feel free to use as many code cells as needed.

        def run_app(img_path):
            ## handle cases for a human face, dog, and neither
```



Sample Human Output

Step 6: Test Your Algorithm
In this section, you will take your new algorithm for a spin! What kind of dog does the algorithm think that *you* look like? If you have a dog, does it predict your dog's breed accurately? If you have a cat, does it mistakenly think that your cat is a dog?

1.1.19 (IMPLEMENTATION) Test Your Algorithm on Sample Images!

Test your algorithm at least six images on your computer. Feel free to use any images you like. Use at least two human and two dog images.

Question 6: Is the output better than you expected :) ? Or worse :(? Provide at least three possible points of improvement for your algorithm.

Answer: (Three possible points for improvement)

```
In [ ]: ## TODO: Execute your algorithm from Step 6 on
        ## at least 6 images on your computer.
        ## Feel free to use as many code cells as needed.

        ## suggested code, below
        for file in np.hstack((human_files[:3], dog_files[:3])):
            run_app(file)
```