

dog_app

January 28, 2021

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 [1]: 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 [2]: 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))
```

```

# 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 [3]: # 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 [13]: 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_count = 0
for img in human_files_short:
    human_count+=int(face_detector(img))
print("Human was detected in "+str(human_count*100/human_files_short.shape[0])+"% of th
dog_count = 0
for img in dog_files_short:
    dog_count+=int(face_detector(img))
print("Human was detected in "+str(dog_count*100/dog_files_short.shape[0])+"% of the do
```

Human was detected in 98.0% of the human images

Human was detected in 17.0% of the dog images

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 [14]: ### (Optional)  
        ### TODO: Test performance of another face 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 [6]: 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 [00:21<00:00, 26284826.87it/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]: from PIL import Image
import torchvision.transforms as transforms
import torchvision.datasets as datasets

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

    image = Image.open(img_path).convert('RGB')

    img_transforms = transforms.Compose([transforms.Resize(224),
                                         transforms.CenterCrop(224),
                                         transforms.ToTensor(),
                                         transforms.Normalize([0.485, 0.456, 0.406],
                                                                [0.229, 0.224, 0.225])])

    # Return the *index* of the predicted class for that image

    transf_img = img_transforms(image)[:3,:,:].unsqueeze(0)

    if use_cuda:
        transf_img = transf_img.cuda()

    output = VGG16.forward(transf_img)
    _, pred = torch.max(output, 1)
    return pred # predicted class index

VGG16_predict(human_files[0])

Out[7]: tensor([ 906])

```

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.
    dog_pred = (VGG16_predict(img_path)>=151)&(VGG16_predict(img_path)<=268)

    return dog_pred # true/false
```

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

human_count = 0
for img in human_files_short:
    human_count+=int(dog_detector(img))
print("Dog was detected in "+str(human_count*100/human_files_short.shape[0])+"% of the h
dog_count = 0
for img in dog_files_short:
    dog_count+=int(dog_detector(img))
print("Dog was detected in "+str(dog_count*100/dog_files_short.shape[0])+"% of the dog i
```

Dog was detected in 2.0% of the human images

Dog was detected in 100.0% of the dog images

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 imbalanced, 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 [11]: import os
         from torchvision import datasets
         from PIL import ImageFile
         ImageFile.LOAD_TRUNCATED_IMAGES = True

         ### TODO: Write data loaders for training, validation, and test sets
         ## Specify appropriate transforms, and batch_sizes
         data_dir = '/data/dog_images'
```



```

# TODO: Define transforms for the training data and testing data
train_transforms = transforms.Compose([transforms.Resize(256),
                                       transforms.CenterCrop(256),
                                       transforms.RandomHorizontalFlip(),
                                       transforms.ToTensor(),
                                       transforms.Normalize([0.485, 0.456, 0.406],
                                                           [0.229, 0.224, 0.225])])

test_transforms = transforms.Compose([transforms.Resize(256),
                                       transforms.CenterCrop(256),
                                       transforms.ToTensor(),
                                       transforms.Normalize([0.485, 0.456, 0.406],
                                                           [0.229, 0.224, 0.225])])

# Pass transforms in here, then run the next cell to see how the transforms look
train_data = datasets.ImageFolder(data_dir + '/train', transform=train_transforms)
val_data = datasets.ImageFolder(data_dir + '/valid', transform=test_transforms)
test_data = datasets.ImageFolder(data_dir + '/test', transform=test_transforms)

loaders_scratch = {}
loaders_scratch['train'] = torch.utils.data.DataLoader(train_data, batch_size=8, shuffle=True)
loaders_scratch['valid'] = torch.utils.data.DataLoader(val_data, batch_size=8)
loaders_scratch['test'] = torch.utils.data.DataLoader(test_data, batch_size=8)

```

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:

My code resize images to 256 by interpolation and also crops it to 256 to standardize. The size was chosen to be 256 so we have a size that is neither too big or too small (and also because it is a multiple of 16 and we will use two pooling layers of 4).

I am augmenting the dataset by a Random horizontal flip so dogs are still in the same vertical direction but we have them in different horizontal directions randomly in each batch.

1.1.8 (IMPLEMENTATION) Model Architecture

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

```

In [12]: 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.C1 = nn.Conv2d(3,64,3,padding=1)
    self.C3 = nn.Conv2d(64,128,3,padding=1)
    self.C5 = nn.Conv2d(128,256,3,padding=1)
    self.pool2 = nn.MaxPool2d(4,4)

    self.L1 = nn.Linear(16*16*256, 500)
    self.L2 = nn.Linear(500, 399)

    self.dropout = nn.Dropout(0.2)

    def forward(self, x):
        ## Define forward behavior
        x = F.relu(self.C1(x))
        x = self.pool2(F.relu(self.C3(x)))
        x = self.pool2(F.relu(self.C5(x)))

        x = x.view(-1,16*16*256)

        # add dropout layer
        x = self.dropout(x)
        # add 1st hidden layer, with relu activation function
        x = F.relu(self.L1(x))
        # add dropout layer
        x = self.dropout(x)
        # add 2nd hidden layer, with relu activation function
        x = self.L2(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
if use_cuda:
    model_scratch.cuda()

```

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

Answer:

I defined 3 convolutional layers that will extract my features and 2 linear layers that will end-up with classification. I tried to get an NN that is deeper enough to extract the features that I need in this very hard classification problem. I added 2 pooling steps to reduce the size of my final image. I also added dropout steps to reduce overfitting and I used relu as my activation function.

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 [13]: import torch.optim as optim

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

        ### TODO: select optimizer
        optimizer_scratch = optim.Adam(model_scratch.parameters(), lr=0.001)
```

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 [14]: def train(n_epochs, loaders, model, optimizer, criterion, use_cuda, save_path):
        """returns trained model"""
        # initialize tracker for minimum validation loss
        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()
                optimizer.zero_grad()
                ## find the loss and update the model parameters accordingly
                output = model(data)
                loss = criterion(output, target)
                loss.backward()
                optimizer.step()
                # record the average training loss, using something like
                train_loss = train_loss + ((1 / (batch_idx + 1)) * (loss.data - 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 + ((1 / (batch_idx + 1)) * (loss.data - 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:
        torch.save(model.state_dict(), save_path)
        valid_loss_min = valid_loss

    # return trained model
    return model

```

```

In [15]: # train the model
         model_scratch = train(10, loaders_scratch, model_scratch, optimizer_scratch,
                               criterion_scratch, use_cuda, 'model_scratch.pt')

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

```

| | | |
|-----------|-------------------------|---------------------------|
| Epoch: 1 | Training Loss: 4.942301 | Validation Loss: 4.691346 |
| Epoch: 2 | Training Loss: 4.449821 | Validation Loss: 4.295545 |
| Epoch: 3 | Training Loss: 4.123222 | Validation Loss: 4.119864 |
| Epoch: 4 | Training Loss: 3.895609 | Validation Loss: 4.086201 |
| Epoch: 5 | Training Loss: 3.618257 | Validation Loss: 4.204495 |
| Epoch: 6 | Training Loss: 3.280044 | Validation Loss: 4.058241 |
| Epoch: 7 | Training Loss: 2.921997 | Validation Loss: 4.298953 |
| Epoch: 8 | Training Loss: 2.572740 | Validation Loss: 4.328243 |
| Epoch: 9 | Training Loss: 2.250106 | Validation Loss: 4.479422 |
| Epoch: 10 | Training Loss: 1.901882 | Validation Loss: 4.706472 |

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 [16]: 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))

```

In [17]: # call test function

```
test(loaders_scratch, model_scratch, criterion_scratch, use_cuda)
```

Test Loss: 4.000412

Test Accuracy: 10% (90/836)

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 [27]: import torch
import os
from torchvision import datasets, transforms
from PIL import ImageFile
ImageFile.LOAD_TRUNCATED_IMAGES = True

### TODO: Write data loaders for training, validation, and test sets
## Specify appropriate transforms, and batch_sizes
data_dir = '/data/dog_images'

# TODO: Define transforms for the training data and testing data
train_transforms = transforms.Compose([transforms.Resize(255),
                                       transforms.CenterCrop(224),
                                       transforms.RandomHorizontalFlip(),
                                       transforms.ToTensor(),
                                       transforms.Normalize([0.485, 0.456, 0.406],
                                                            [0.229, 0.224, 0.225])])

test_transforms = transforms.Compose([transforms.Resize(255),
                                       transforms.CenterCrop(224),
                                       transforms.ToTensor(),
                                       transforms.Normalize([0.485, 0.456, 0.406],
                                                            [0.229, 0.224, 0.225])])

# Pass transforms in here, then run the next cell to see how the transforms look
train_data = datasets.ImageFolder(data_dir + '/train', transform=train_transforms)
val_data = datasets.ImageFolder(data_dir + '/valid', transform=test_transforms)
test_data = datasets.ImageFolder(data_dir + '/test', transform=test_transforms)

loaders_transfer = {}
loaders_transfer['train'] = torch.utils.data.DataLoader(train_data, batch_size=8, shuffle=True)
loaders_transfer['valid'] = torch.utils.data.DataLoader(val_data, batch_size=8)
loaders_transfer['test'] = torch.utils.data.DataLoader(test_data, batch_size=8)

```

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 [29]: import torchvision.models as models
import torch.nn as nn

## TODO: Specify model architecture
model_transfer = models.vgg16(pretrained=True)

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

```

```
model_transfer.classifier[6] = nn.Linear(4096,399)
```

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

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: As we have seen before, this vgg16 is trained on ImageNet, which has dog breeds just as we want to have so maybe we can rely on the feature part of the network to identify the breeds. So I have frozen the feature part of the network and I replaced the final classifier node by one that has the final number of classes (breeds) required.

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 [20]: from torch import optim
```

```
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 [21]: import numpy as np
```

```
# train the model
model_transfer = train(10, loaders_transfer, model_transfer, optimizer_transfer, criterion_transfer)

# 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: 2.609741 | Validation Loss: 0.754379 |
| Epoch: 2 | Training Loss: 0.789838 | Validation Loss: 0.537581 |
| Epoch: 3 | Training Loss: 0.562883 | Validation Loss: 0.458209 |
| Epoch: 4 | Training Loss: 0.457998 | Validation Loss: 0.428795 |
| Epoch: 5 | Training Loss: 0.388864 | Validation Loss: 0.399844 |
| Epoch: 6 | Training Loss: 0.331347 | Validation Loss: 0.389722 |
| Epoch: 7 | Training Loss: 0.306690 | Validation Loss: 0.383744 |
| Epoch: 8 | Training Loss: 0.273008 | Validation Loss: 0.371585 |
| Epoch: 9 | Training Loss: 0.248607 | Validation Loss: 0.381316 |
| Epoch: 10 | Training Loss: 0.222697 | Validation Loss: 0.359524 |

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 [22]: test(loaders_transfer, model_transfer, criterion_transfer, use_cuda)
```

Test Loss: 0.399148

Test Accuracy: 87% (734/836)

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 [28]: ### TODO: Write a function that takes a path to an image as input
### and returns the dog breed that is predicted by the model.
from PIL import Image
import torchvision.models as models
import numpy as np
from glob import glob
import torch.nn as nn

# 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 train_data.classes]
dog_files = np.array(glob("/data/dog_images/**/*.jpg"))

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

    image = Image.open(img_path).convert('RGB')

    img_transforms = transforms.Compose([transforms.Resize(255),
                                         transforms.CenterCrop(224),
                                         transforms.ToTensor(),
                                         transforms.Normalize([0.485, 0.456, 0.406],
                                                                [0.229, 0.224, 0.225])])

    # Return the *index* of the predicted class for that image

    transf_img = img_transforms(image)[:3, :, :].unsqueeze(0)

    ## TODO: Specify model architecture
    model_transfer = models.vgg16()
    model_transfer.classifier[6] = nn.Linear(4096, 399)
    model_transfer = model_transfer.cuda()
    model_transfer.load_state_dict(torch.load('model_transfer.pt'))
```




Sample Human Output

```
model_transfer.eval()

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

# move model to GPU if CUDA is available
if use_cuda:
    transf_img = transf_img.cuda()
else:
    model_transfer.device('cpu')

output = model_transfer.forward(transf_img)
_, pred = torch.max(output, 1)

return class_names[pred]

predict_breed_transfer(dog_files[0])
```

Out[28]: 'Mastiff'

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

```

import cv2

def run_app(img_path):
    if dog_detector(img_path):
        print("Hello Dog!")
        img = cv2.imread(img_path)

        plt.imshow(img)
        plt.show()

        print("Your predicted breed is ..")
        print(predict_breed_transfer(img_path))
    elif face_detector(img_path):
        print("Hello Human!")
        img = cv2.imread(img_path)

        plt.imshow(img)
        plt.show()

        print("You look like a..")
        print(predict_breed_transfer(img_path))
    else:
        print("Error! Give me a human face or dog")

```

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: I really like the output but three possible points of improvement for the neural network that detects dog breeds could be: - to randomize the final feature layer of vgg16 so it can better capture the dog features - to add more linear layers to our classification - to augment the data even further for retraining the vgg16

For the algorithm: - directly use the vgg16 trained with ImageNet to detect the breeds for the dog detector instead of having two different networks to do almost the same job - retrain the vgg16 leveraging the breeds and adding a class for human. But this would not allow us to detect dog breeds for humans, it is just an idea.

```

In [32]: ## TODO: Execute your algorithm from Step 6 on
         ## at least 6 images on your computer.

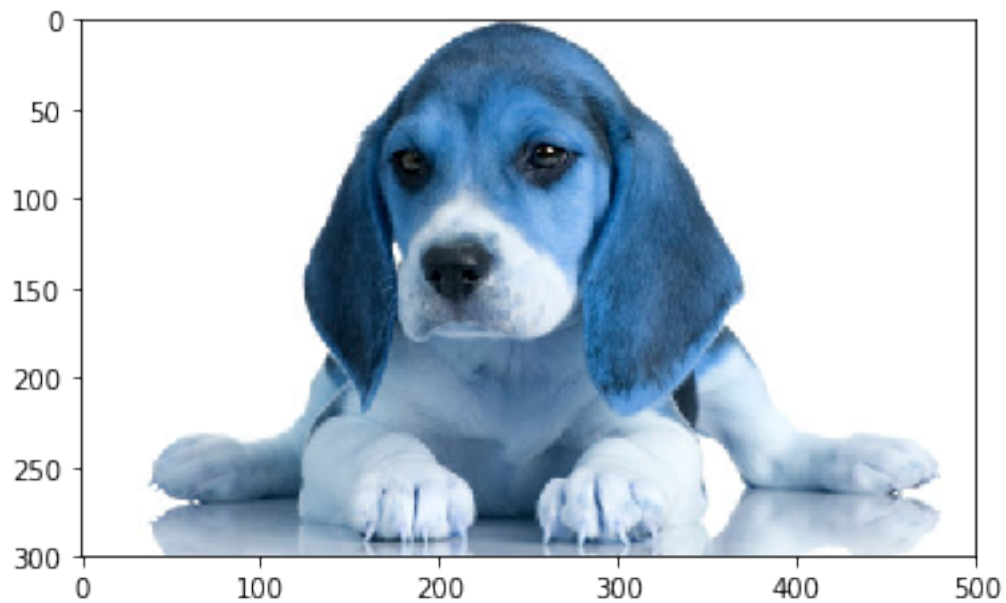
```

```
## Feel free to use as many code cells as needed.
```

```
## suggested code, below
```

```
for file in glob("imgs/*"):  
    run_app(file)
```

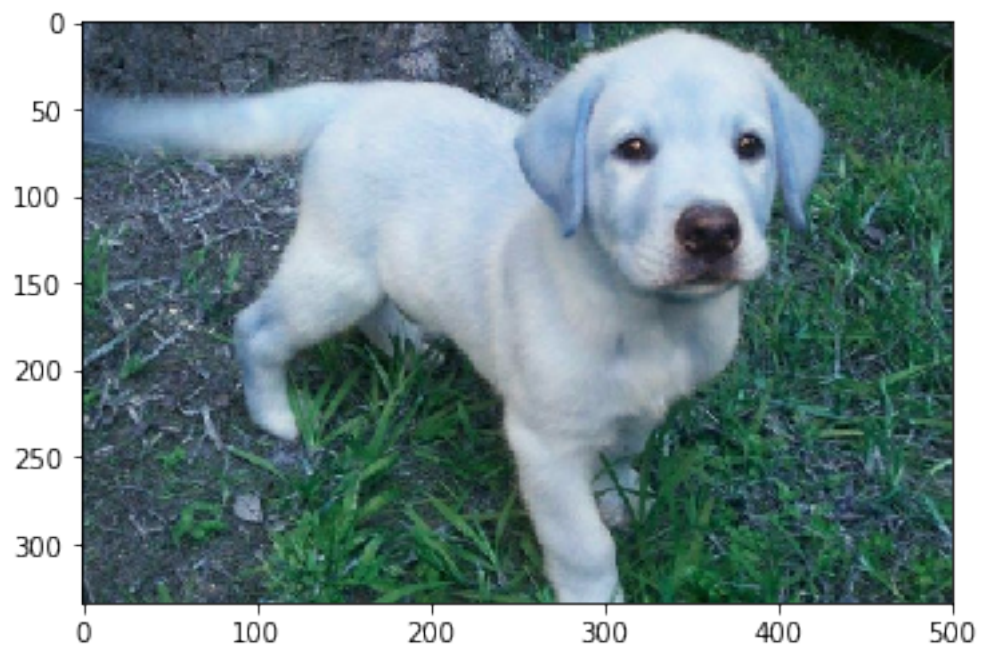
Hello Dog!



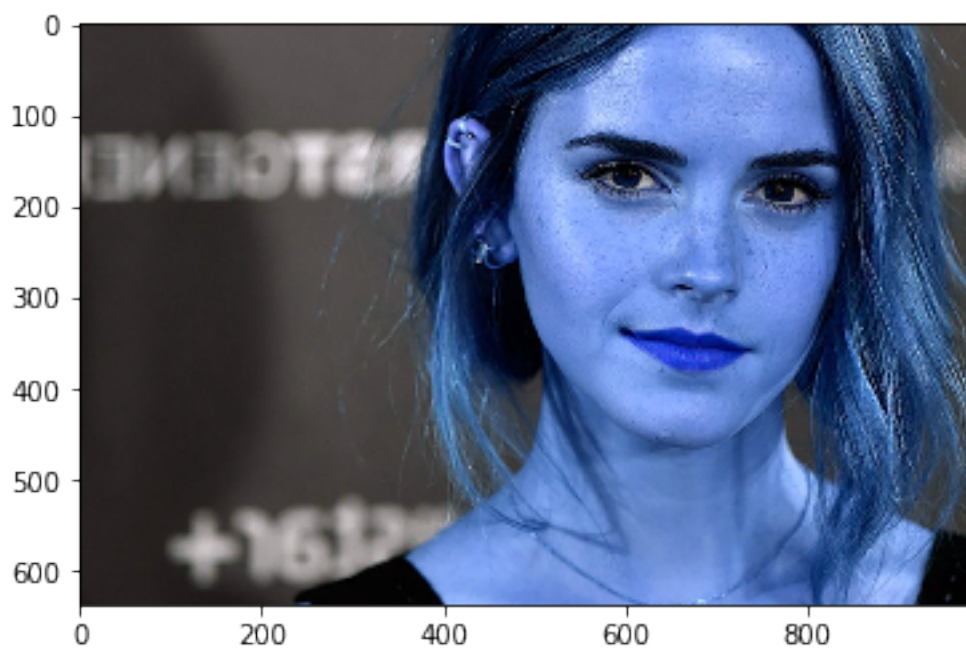
Your predicted breed is ..

Beagle

Hello Dog!



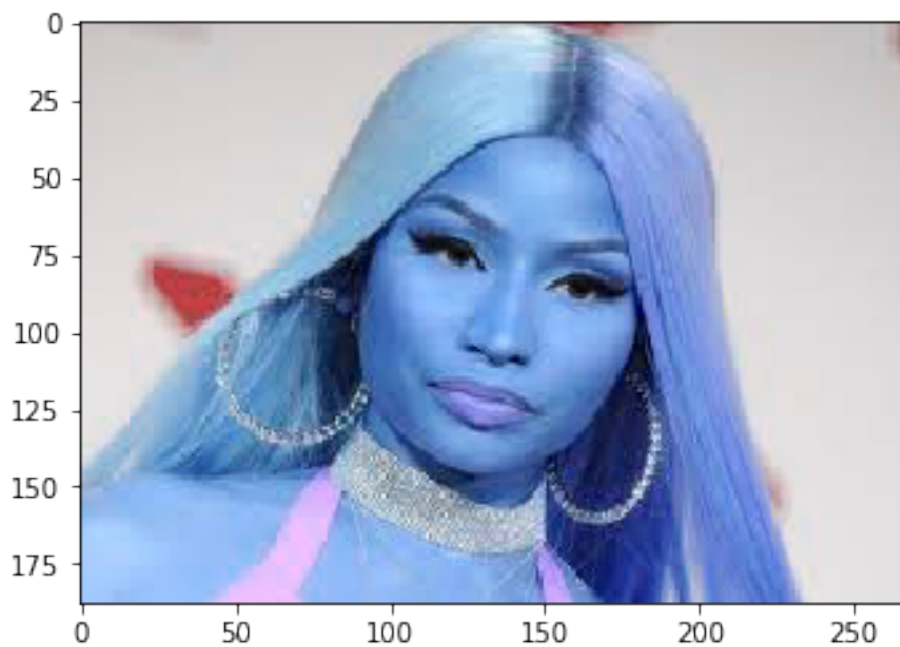
Your predicted breed is ..
Labrador retriever
Hello Human!



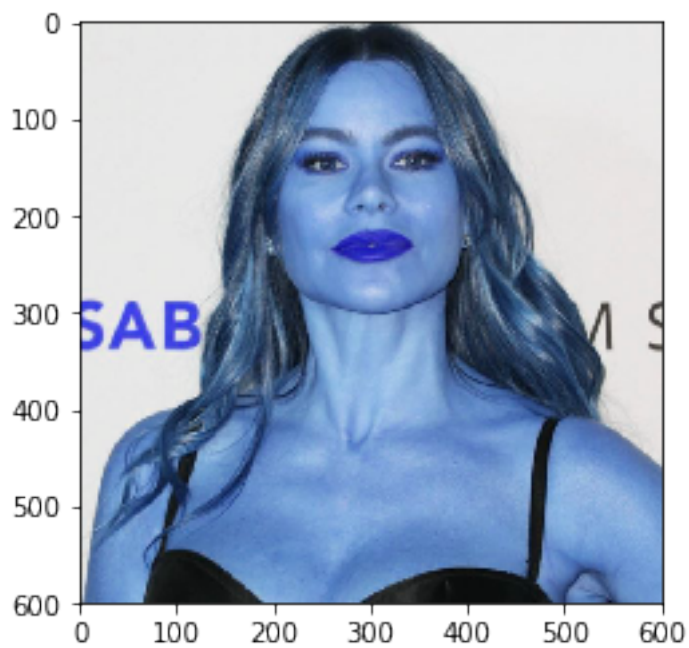
You look like a..
Nova scotia duck tolling retriever
Error! Give me a human face or dog
Hello Dog!



Your predicted breed is ..
Bichon frise
Hello Human!



You look like a..
Dachshund
Hello Human!



```
You look like a..  
Chinese crested
```

```
In [ ]:
```