

dog_app

June 9, 2020

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 [62]: 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 [63]: 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 [64]: # 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:

Number of human faces detected correctly in `human_files` are: 98 with percentage of: 98.0 %.

Number of human faces detected wrongly in `dog_files` are: 17 with percentage of: 17.0 %.

```
In [66]: 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.
HumanFacesInHumanFiles = 0
HumanFacesInDogFiles = 0
for humanPic in human_files_short:
    if face_detector(humanPic):
        HumanFacesInHumanFiles = HumanFacesInHumanFiles + 1

for dogPic in dog_files_short:
    if face_detector(dogPic):
        HumanFacesInDogFiles = HumanFacesInDogFiles + 1

print("Number of human faces detected correctly in human_files are: " , HumanFacesInHumanFiles)
print("Number of human faces detected wrongly in dog_files are: " , HumanFacesInDogFiles)
```

Number of human faces detected correctly in human_files are: 98 with percentage of: 98.0 %.
Number of human faces detected wrongly in dog_files are: 17 with percentage of: 17.0 %.

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 [67]: ### (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 [68]: 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()
```

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 [69]: from PIL import Image
import torchvision.transforms as transforms

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

    normalize = transforms.Normalize( mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.
    # VGG-16 Takes 224x224 images as input, so we resize all of them and convert data t
    transform = transforms.Compose([transforms.Resize(256), transforms.CenterCrop(224),

    # add the batch dimension, models want batch
    image = transform(image)[:3,:,:].unsqueeze(0)
    if use_cuda:
        image = image.cuda()
    predict = VGG16(image)
    predict = predict.data.cpu().argmax()

    return predict # 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 [70]: ### returns "True" if a dog is detected in the image stored at img_path
def dog_detector(img_path):
    ## TODO: Complete the function.
    classIndex = VGG16_predict(img_path)

```

```
return ( (classIndex >= 151) & ( classIndex <= 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:

Number of dogs detected correctly in `human_files` are: 1 with percentage of: 1.0 %.

Number of human faces detected wrongly in `dog_files` are: 100 with percentage of: 100.0 %.

```
In [71]: ### TODO: Test the performance of the dog_detector function
### on the images in human_files_short and dog_files_short.
DogsInHumanFiles = 0
DogsInDogFiles   = 0
for humanPic in human_files_short:
    if dog_detector(humanPic):
        DogsInHumanFiles = DogsInHumanFiles + 1

for dogPic in dog_files_short:
    if dog_detector(dogPic):
        DogsInDogFiles = DogsInDogFiles + 1

print("Number of dogs detected correctly in human_files are: " , DogsInHumanFiles, " wi
print("Number of human faces detected wrongly in dog_files are: " , DogsInDogFiles, " w
```

Number of dogs detected correctly in `human_files` are: 1 with percentage of: 1.0 %.

Number of human faces detected wrongly in `dog_files` are: 100 with percentage of: 100.0 %.

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 [72]: ### (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 [73]: import os
         from torchvision import datasets

         ### TODO: Write data loaders for training, validation, and test sets
         ## Specify appropriate transforms, and batch_sizes
         from PIL import ImageFile
         ImageFile.LOAD_TRUNCATED_IMAGES = True
         # number of subprocesses to use for data loading
         num_workers = 0
         # how many samples per batch to load
         batch_size = 20

         data_dir = '/data/dog_images/'
         train_dir = os.path.join(data_dir, 'train/')
         test_dir = os.path.join(data_dir, 'test/')
```

```

validation_dir = os.path.join(data_dir, 'valid/')

#Augmentation "rotation , flip" will be applied only on training data
data_transforms = {
    'train' : transforms.Compose([transforms.Resize(224),
                                  transforms.CenterCrop(224),
                                  transforms.RandomHorizontalFlip(), # randomly flip an
                                  transforms.RandomRotation(10),
                                  transforms.ToTensor(),
                                  transforms.Normalize(mean=[0.485, 0.456, 0.406],
                                                         std=[0.229, 0.224, 0.225]))),

    'test' : transforms.Compose([transforms.Resize(224),
                                  transforms.CenterCrop(224),
                                  transforms.ToTensor(),
                                  transforms.Normalize(mean=[0.485, 0.456, 0.406],
                                                         std=[0.229, 0.224, 0.225]))),

    'validation' : transforms.Compose([transforms.Resize(224),
                                       transforms.CenterCrop(224),
                                       transforms.ToTensor(),
                                       transforms.Normalize(mean=[0.485, 0.456, 0.406],
                                                              std=[0.229, 0.224, 0.225]))])

imageDatasets = {'train' : datasets.ImageFolder(root=train_dir,transform=data_transforms),
                 'test' : datasets.ImageFolder(root=test_dir,transform=data_transforms),
                 'validation' : datasets.ImageFolder(root=validation_dir,transform=data_transforms)}

loaders_scratch = {'train' : torch.utils.data.DataLoader(imageDatasets['train'],batch_size=16),
                   'test' : torch.utils.data.DataLoader(imageDatasets['test'],batch_size=16),
                   'validation' : torch.utils.data.DataLoader(imageDatasets['validation'],batch_size=16)}

```

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: - I decided to resize the image to 244x244 using transforms.Resize(224) and center and normalize it, as i found Imagenet requires this size and also to reduce memory consumption and training time. and then transform it to tensors that can be passed to our network. - Yes i decided to augment the training dataset by random flip using transforms.RandomHorizontalFlip() and images rotation using transforms.RandomRotation(10)

1.1.8 (IMPLEMENTATION) Model Architecture

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

```
In [74]: 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
        # convolutional layer (sees 32x32x3 image tensor)
        self.conv1 = nn.Conv2d(3, 16, 3, padding=1)
        # convolutional layer (sees 16x16x16 tensor)
        self.conv2 = nn.Conv2d(16, 32, 3, padding=1)
        # convolutional layer (sees 8x8x32 tensor)
        self.conv3 = nn.Conv2d(32, 64, 3, padding=1)
        # max pooling layer
        self.pool = nn.MaxPool2d(2, 2)
        # linear layer (64 * 28 * 28 -> 500)
        self.fc1 = nn.Linear(64 * 28 * 28, 500)
        # linear layer (500 -> 133)
        self.fc2 = nn.Linear(500, 133)
        # dropout layer (p=0.25)
        self.dropout = nn.Dropout(0.25)
    def forward(self, x):
        ## Define forward behavior
        x = self.pool(F.relu(self.conv1(x)))
        x = self.pool(F.relu(self.conv2(x)))
        x = self.pool(F.relu(self.conv3(x)))
        # flatten image input
        x = x.view(-1, 64 * 28 * 28)
        # add dropout layer
        x = self.dropout(x)
        # add 1st hidden layer, with relu activation function
        x = F.relu(self.fc1(x))
        # add dropout layer
        x = self.dropout(x)
        # add 2nd hidden layer, with relu activation function
        x = self.fc2(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()

print(model_scratch)

```

```

Net(
  (conv1): Conv2d(3, 16, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (conv2): Conv2d(16, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (conv3): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (fc1): Linear(in_features=50176, out_features=500, bias=True)
  (fc2): Linear(in_features=500, out_features=133, bias=True)
  (dropout): Dropout(p=0.25)
)

```

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

Answer:

- Let's start with the input: since the input image is RGB and 224x224 pixels >> so we will have our input 3x224x224
- Finally since we have 133 classes in our image datasets so our final layer will have 133 outputs
- In between i created 3 convolutional layers with relu activation function and max pooling layer (2,2) which reduce the x-y size of an input to half, keeping only the most active pixels from the previous layer.
- The usual Linear + Dropout layers to avoid overfitting and produce a 113-dim output
- For the details:

```

(conv1): Conv2d(3, 16, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(conv2): Conv2d(16, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(conv3): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
(fc1): Linear(in_features=50176, out_features=500, bias=True)
(fc2): Linear(in_features=500, out_features=133, bias=True)
(dropout): Dropout(p=0.25)

```

(conv1) (3,224,224) as input with depth = 3 and converted it into a depth of 16 layers, the filter used was 3x3 with stride of 1 and padding of 1 to avoid removing edges and so on till conv3 with depth 64

afterwards, 2 full connected Linear layers with relu activation function were added, dropout is 25% to avoid overfitting problem first linear layer(fc1): (64 28 28 , 500). second linear layer(fc2): (500 , 133).

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 [75]: import torch.optim as optim
```

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

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

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 [76]: 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()
            ## 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_loss))
            # clear the gradients of all optimized variables
            optimizer.zero_grad()
            # forward pass: compute predicted outputs by passing inputs to the model
            output = model(data)
            # calculate the batch loss
            loss = criterion(output, target)
            # backward pass: compute gradient of the loss with respect to model parameters
            loss.backward()
            # perform a single optimization step (parameter update)
            optimizer.step()
            # update training loss
            ## 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['validation']):
    # move to GPU
    if use_cuda:
        data, target = data.cuda(), target.cuda()
    ## update the average validation loss
    # forward pass: compute predicted outputs by passing inputs to the model
    output = model(data)
    # calculate the batch loss
    loss = criterion(output, target)
    # update average validation loss
    valid_loss = valid_loss + ((1 / (batch_idx + 1)) * (loss.data - valid_loss))

# calculate average losses
train_loss = train_loss/len(imageDatasets['train'])
valid_loss = valid_loss/len(imageDatasets['validation'])

# 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)

    print('Validation loss decreased ({:.6f} --> {:.6f}). Saving model ...'.format(
        valid_loss_min,
        valid_loss))

    valid_loss_min = valid_loss
# return trained model
return model

# train the model
model_scratch = train(12, loaders_scratch, model_scratch, optimizer_scratch, criterion_

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

```

```

Epoch: 1          Training Loss: 0.000730          Validation Loss: 0.005802
Validation loss decreased (inf --> 0.005802). Saving model ...
Epoch: 2          Training Loss: 0.000713          Validation Loss: 0.005582

```

```

Validation loss decreased (0.005802 --> 0.005582). Saving model ...
Epoch: 3      Training Loss: 0.000690      Validation Loss: 0.005463
Validation loss decreased (0.005582 --> 0.005463). Saving model ...
Epoch: 4      Training Loss: 0.000671      Validation Loss: 0.005346
Validation loss decreased (0.005463 --> 0.005346). Saving model ...
Epoch: 5      Training Loss: 0.000649      Validation Loss: 0.005218
Validation loss decreased (0.005346 --> 0.005218). Saving model ...
Epoch: 6      Training Loss: 0.000630      Validation Loss: 0.005120
Validation loss decreased (0.005218 --> 0.005120). Saving model ...
Epoch: 7      Training Loss: 0.000615      Validation Loss: 0.005089
Validation loss decreased (0.005120 --> 0.005089). Saving model ...
Epoch: 8      Training Loss: 0.000600      Validation Loss: 0.005064
Validation loss decreased (0.005089 --> 0.005064). Saving model ...
Epoch: 9      Training Loss: 0.000587      Validation Loss: 0.004922
Validation loss decreased (0.005064 --> 0.004922). Saving model ...
Epoch: 10     Training Loss: 0.000571      Validation Loss: 0.004999
Epoch: 11     Training Loss: 0.000555      Validation Loss: 0.004902
Validation loss decreased (0.004922 --> 0.004902). Saving model ...
Epoch: 12     Training Loss: 0.000539      Validation Loss: 0.004915

```

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 [77]: 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: 4.081820

Test Accuracy: 10% (89/836)

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 [18]: ## TODO: Specify data loaders
         # Will use the same loaders

```

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

         ## TODO: Specify model architecture
         # Load the pretrained model from pytorch
         model_transfer = models.vgg16(pretrained=True)
         # print out the model structure
         print(model_transfer)

         # Freeze training for all features layers
         for param in model_transfer.features.parameters():

```



```

        param.requires_grad = False

    n_inputs = model_transfer.classifier[6].in_features

    # add last linear layer (n_inputs , 133 classes)
    last_layer = nn.Linear(n_inputs, 133)

    model_transfer.classifier[6] = last_layer

    if use_cuda:
        model_transfer = model_transfer.cuda()

VGG(
(features): Sequential(
  (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (1): ReLU(inplace)
  (2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (3): ReLU(inplace)
  (4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (5): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (6): ReLU(inplace)
  (7): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (8): ReLU(inplace)
  (9): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (10): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (11): ReLU(inplace)
  (12): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (13): ReLU(inplace)
  (14): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (15): ReLU(inplace)
  (16): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (17): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (18): ReLU(inplace)
  (19): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (20): ReLU(inplace)
  (21): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (22): ReLU(inplace)
  (23): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (24): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (25): ReLU(inplace)
  (26): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (27): ReLU(inplace)
  (28): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (29): ReLU(inplace)
  (30): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
)
(classifier): Sequential(
  (0): Linear(in_features=25088, out_features=4096, bias=True)

```

```

(1): ReLU(inplace)
(2): Dropout(p=0.5)
(3): Linear(in_features=4096, out_features=4096, bias=True)
(4): ReLU(inplace)
(5): Dropout(p=0.5)
(6): Linear(in_features=4096, out_features=1000, bias=True)
)
)

```

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:

I chose VGG16 model to be used in transfer learning. I freeze training for all "features" layers to take advantage of its training. I changed the last linear layer output to be 133 classes to fit with our output.

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]: 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 [ ]: # train the model
        model_transfer = train(12, loaders_scratch, model_transfer, optimizer_transfer, criterion_transfer)
        # train(n_epochs, 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: 0.000060      Validation Loss: 0.000462
Validation loss decreased (inf --> 0.000462). Saving model ...
Epoch: 2      Training Loss: 0.000057      Validation Loss: 0.000461
Validation loss decreased (0.000462 --> 0.000461). Saving model ...
Epoch: 3      Training Loss: 0.000055      Validation Loss: 0.000454
Validation loss decreased (0.000461 --> 0.000454). Saving model ...
Epoch: 4      Training Loss: 0.000054      Validation Loss: 0.000444
Validation loss decreased (0.000454 --> 0.000444). Saving model ...
Epoch: 5      Training Loss: 0.000050      Validation Loss: 0.000448
Epoch: 6      Training Loss: 0.000048      Validation Loss: 0.000438
Validation loss decreased (0.000444 --> 0.000438). Saving model ...
Epoch: 7      Training Loss: 0.000047      Validation Loss: 0.000441

```

```
Epoch: 8           Training Loss: 0.000045           Validation Loss: 0.000426
Validation loss decreased (0.000438 --> 0.000426). Saving model ...
Epoch: 9           Training Loss: 0.000043           Validation Loss: 0.000434
Epoch: 10          Training Loss: 0.000042           Validation Loss: 0.000430
Epoch: 11          Training Loss: 0.000043           Validation Loss: 0.000421
Validation loss decreased (0.000426 --> 0.000421). Saving model ...
```

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 [30]: test(loaders_scratch, model_transfer, criterion_transfer, use_cuda)
```

```
Test Loss: 0.440692
```

```
Test Accuracy: 86% (725/836)
```

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 [78]: ### 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 imageDatasets['train'].classes]


def predict_breed_transfer(img_path):
    # load the image and return the predicted breed
    img=Image.open(img_path)
    transform=transforms.Compose([ transforms.Resize((224,224)), transforms.ToTensor()],
    #from image to tensor
    image_tensor = transform(img).float()
    #adds a dimension with a length of one
    imageTen = image_tensor.unsqueeze_(0)
    if use_cuda:
        image = imageTen.cuda()
    predict = model_transfer(image)
    predict = predict.data.cpu().argmax()

    return class_names[predict]
```

```

hello, human!
0
200
400
600
800
1000
1200
1400
0 500 1000
You look like a ...
Chinese_shar-pei

```



The image shows a woman's face with a bounding box. The text 'hello, human!' is at the top, and 'You look like a ... Chinese_shar-pei' is at the bottom. The bounding box is a rectangle around the woman's face, with a vertical axis on the left and a horizontal axis at the bottom.

Sample Human Output

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 [79]: ### TODO: Write your algorithm.
         ### Feel free to use as many code cells as needed.
def display_img(img_path):
    img = Image.open(img_path)
    plt.imshow(img)
    plt.show()

def run_app(img_path):
    ## handle cases for a human face, dog, and neither
    predicted_breed = predict_breed_transfer (img_path)
    if dog_detector(img_path):
        print ('\n\n hello, it is a dog ')
        display_img(img_path)
        return print ('the predicted breed is:', predicted_breed)

    elif face_detector(img_path):
        print ('\n\n hello, human')
        display_img(img_path)
        return print (' You look like a ', predicted_breed)

    else:

```

```
display_img(img_path)
print ('\n\n error: sorry it is neither human, nor 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.

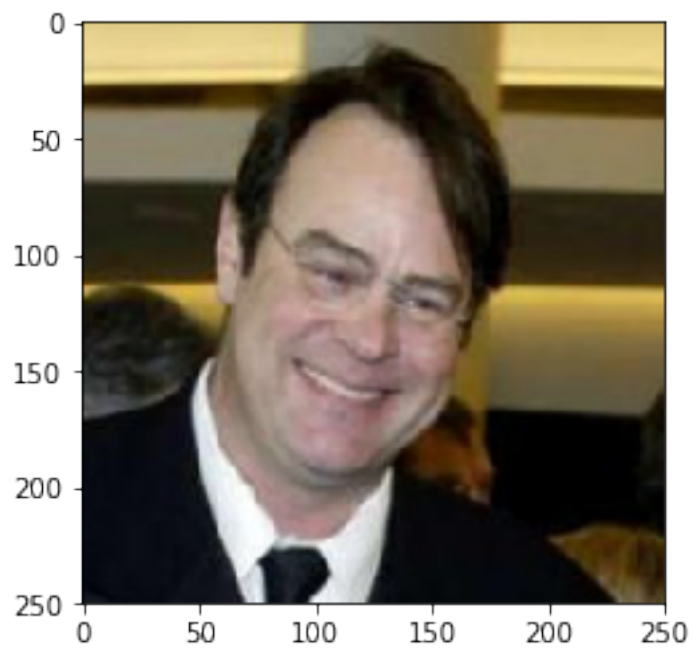
Answer: (Three possible points for improvement)

- 1- if the model is trained with larger number of epochs, it may be better.
- 2- train on larger number of human images
- 3- if we choose another architecture it may be faster in prediction with same or better accuracy.
- 4- apply more augmentation
- 5- we can use k fold cross validation when comparing models

```
In [80]: ## 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.

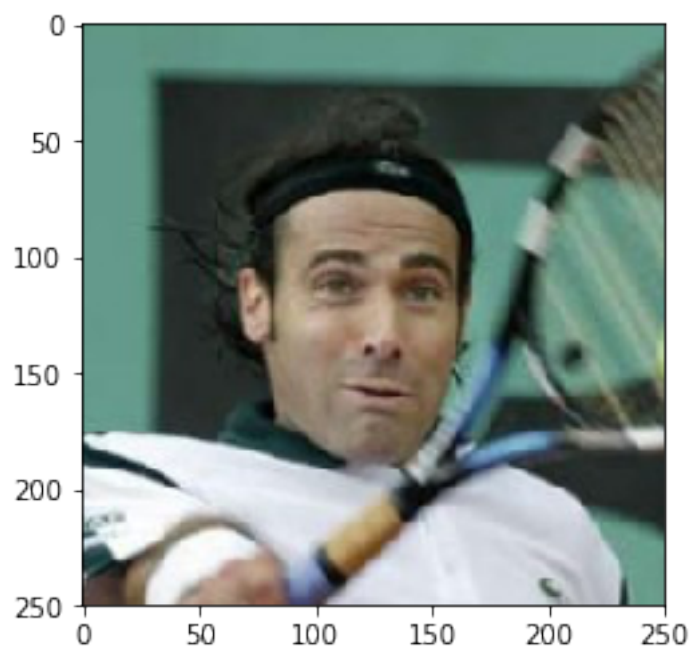
         ## suggested code, below
         for file in np.hstack((human_files[:3], dog_files[:3])):
             run_app(file)
```

hello, human



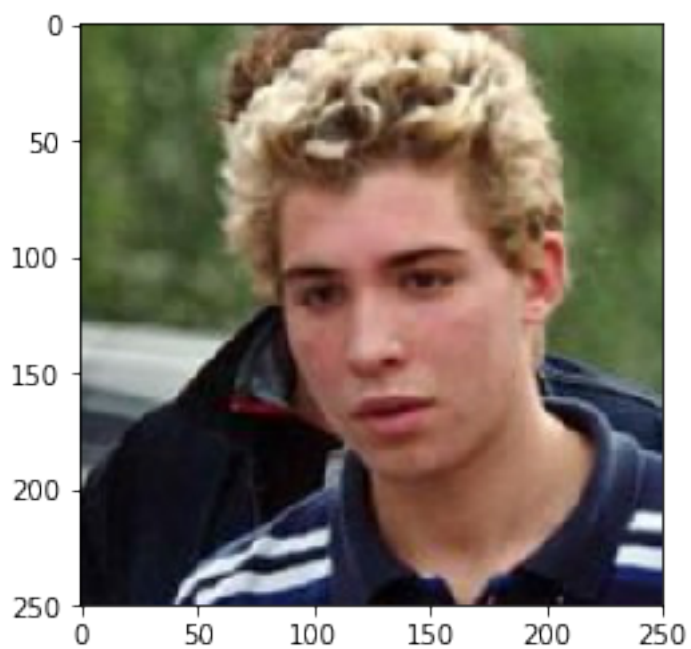
You look like a Brittany

hello, human



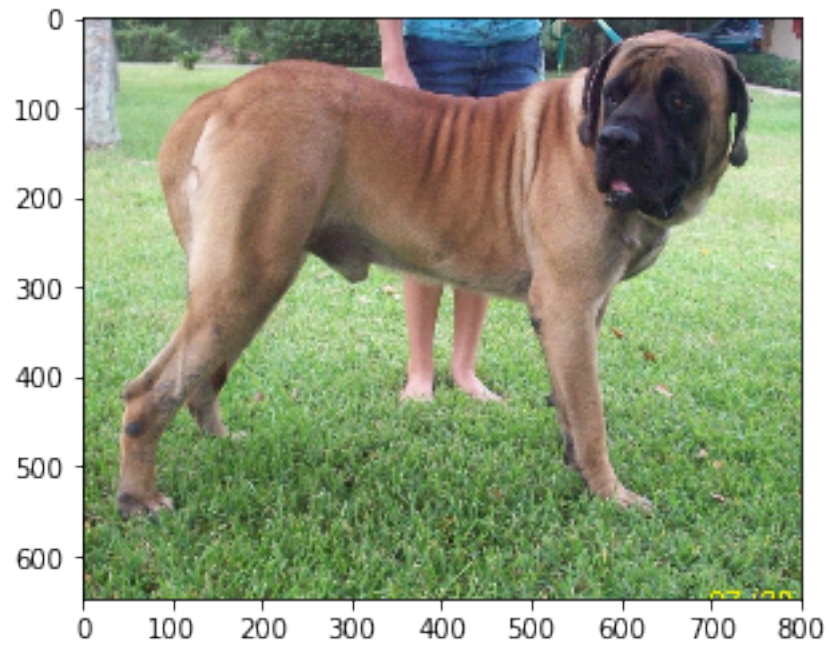
You look like a Dachshund

hello, human



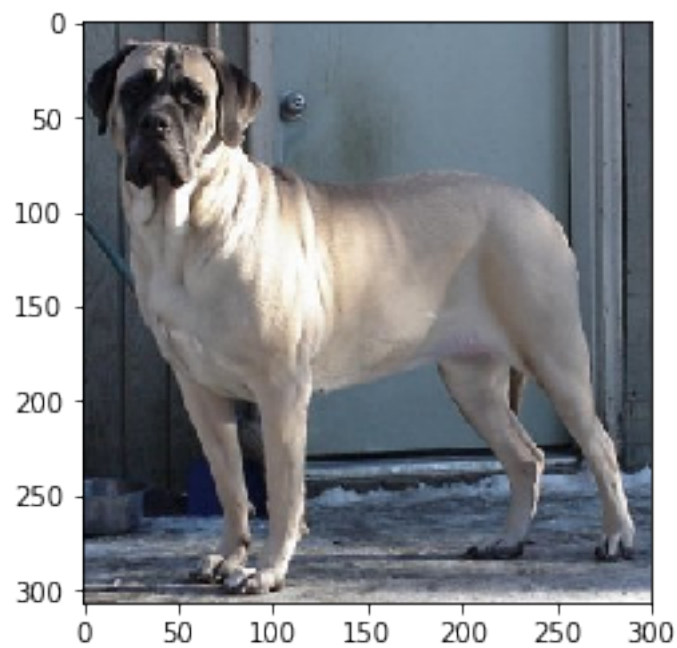
You look like a Portuguese water dog

hello, it is a dog



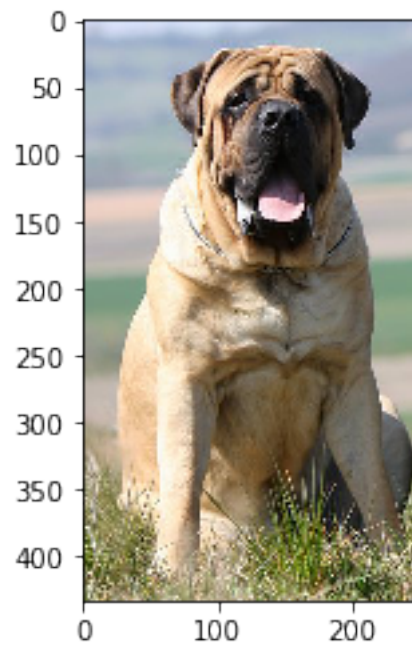
the predicted breed is: Bullmastiff

hello, it is a dog



the predicted breed is: Mastiff

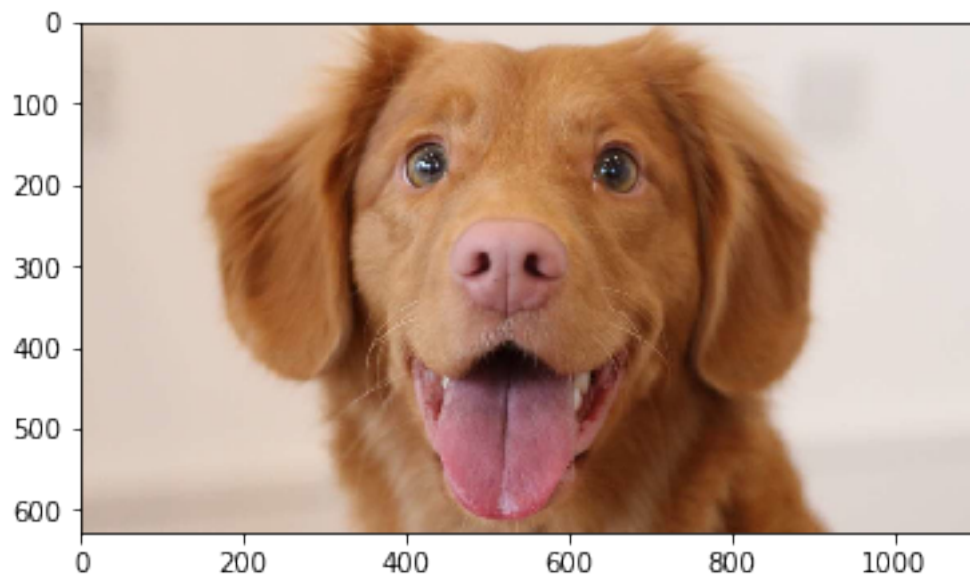
hello, it is a dog



the predicted breed is: Bullmastiff

```
In [81]: test_files = np.array(glob("./test/*"))
         for file in np.hstack(test_files):
             run_app(file)
```

hello, it is a dog



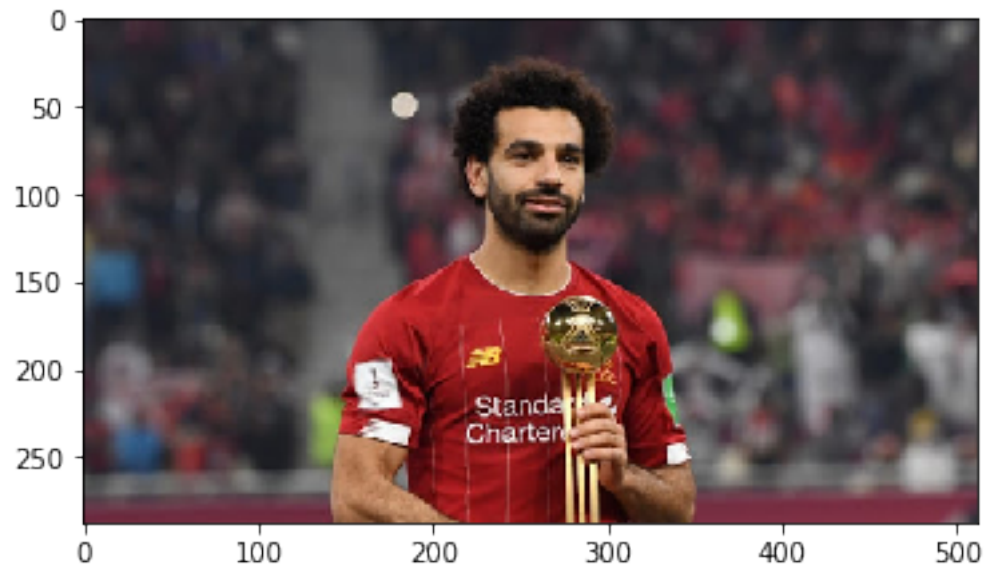
the predicted breed is: Nova scotia duck tolling retriever

hello, it is a dog



the predicted breed is: Golden retriever

hello, human

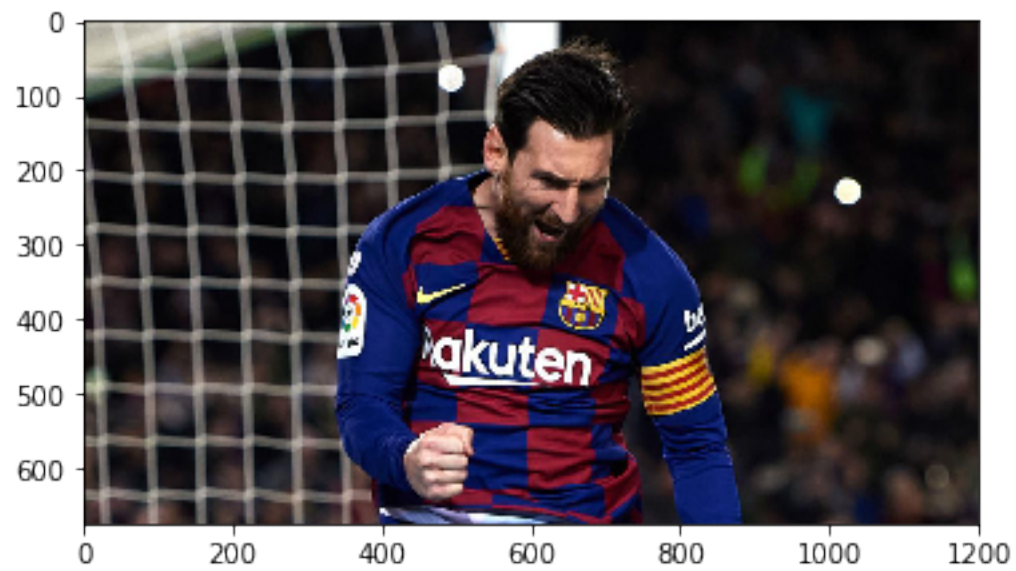


You look like a Pharaoh hound

hello, it is a dog

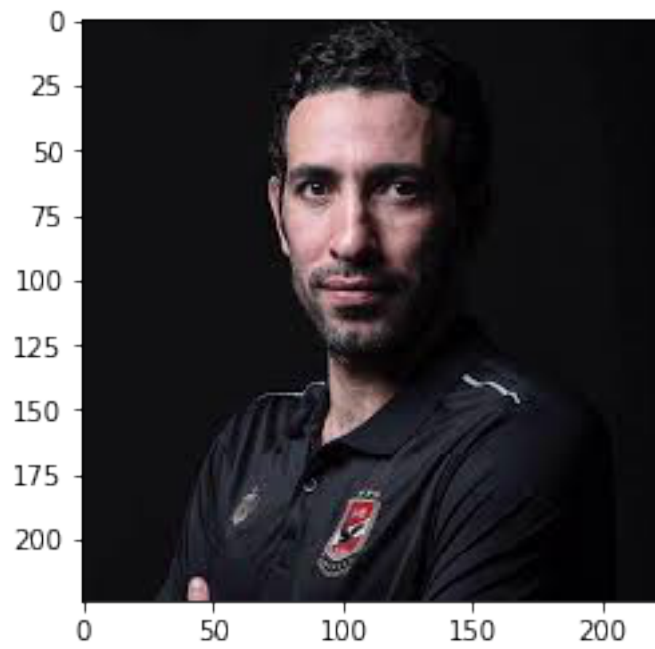


the predicted breed is: Pomeranian



error: sorry it is neither human, nor dog

hello, human



You look like a Maltese

In []: