dog_app

February 20, 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

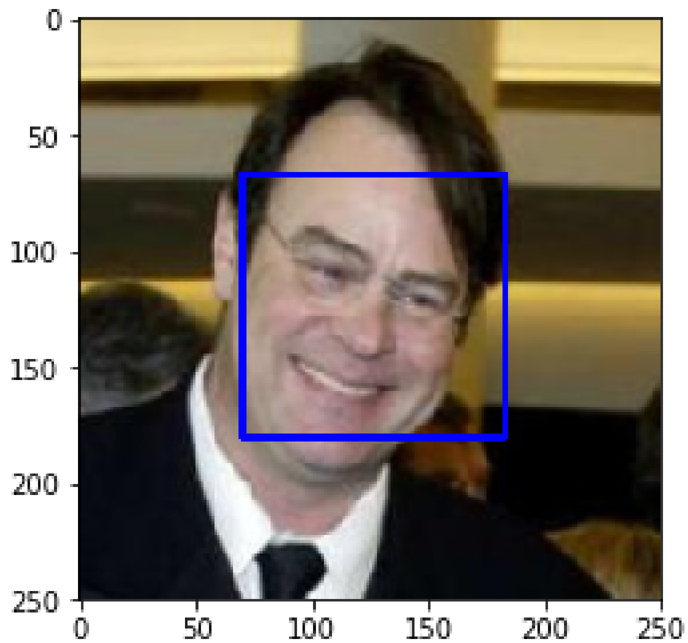
In this section, we use OpenCV's implementation of Haar feature-based cascade classifiers to detect human faces in images.

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))
```

```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 [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: - 98% of the First 100 images in `human_files` detected human face - 17% of the First 100 images in `dog_files` detected human face

```
In [4]: 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_acc = 0.0
        dog_acc = 0.0
        for human, dog in zip(human_files_short, dog_files_short):
            if(face_detector(human)):
                human_acc += 1
            if(face_detector(dog)):
                dog_acc += 1
        else:
            print('Human Accuracy is {:.2f}% \t Dog Accuracy is {:.2f}%'
                .format(
                    human_acc/len(human_files_short) * 100,
                    dog_acc/len(dog_files_short) * 100
                 )
            )
```

```
Human Accuracy is 98.00%         Dog Accuracy is 17.00%
```

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.

        def convertToRGB(img):
            return cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

        #load cascade classifier training file for lbpcascade
        lbp_face_cascade = cv2.CascadeClassifier('lbpcascades/lbpcascade_frontalface_improved.xm

        def lbp_face_detector(img_path):
            img = cv2.imread(img_path)
            gray = convertToRGB(img)
            faces = lbp_face_cascade.detectMultiScale(gray)
            return len(faces) > 0
```

```
In [6]: lbp_human_acc = 0.0
        lbp_dog_acc = 0.0
        for human, dog in zip(human_files_short, dog_files_short):
            if(lbp_face_detector(human)):
                lbp_human_acc += 1
            if(lbp_face_detector(dog)):
                lbp_dog_acc += 1
        else:
            print('Human Accuracy (LBP) is {:.2f}% \t Dog Accuracy (LBP) is {:.2f}%'
                .format(
                    lbp_human_acc/len(human_files_short) * 100,
                    lbp_dog_acc/len(dog_files_short) * 100
                 )
            )
```

```
Human Accuracy (LBP) is 82.00%         Dog Accuracy (LBP) is 7.00%
```

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 [7]:
```
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()
VGG16
```

Out[7]:
```
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)
  )
)
```

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 [8]:
```
from PIL import Image
import torchvision.transforms as transforms

def image_to_tensor(img_path):
    image = Image.open(img_path).convert('RGB')
    tranform = transforms.Compose([transforms.Resize(size=224),
                                    transforms.CenterCrop((224,224)),
                                   transforms.ToTensor(),
                                   transforms.Normalize(mean=[0.485, 0.456, 0.406]
                                                        std=[0.229, 0.224, 0.225])
    image_tensor = tranform(image)[:3,:,:].unsqueeze(0)
    return image_tensor


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_tensor = image_to_tensor(img_path)

        VGG16.eval() # ------------- put model in evaluation mode ----------------

        # if GPU available, move model inputs to GPU
        if use_cuda:
            image_tensor = image_tensor.cuda()
        output = VGG16(image_tensor)
        _, pred_k = output.topk(1, dim=1)

        pred_k = np.squeeze(pred_k.numpy()) if not use_cuda else np.squeeze(pred_k.cpu().num

        return pred_k # 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 [9]: ### returns "True" if a dog is detected in the image stored at img_path
        def dog_detector(img_path):
            ## TODO: Complete the function.
            pred = VGG16_predict(img_path)
            if pred >= 151 and pred <= 268:
                return True
            else:
                return 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: - Human Accuracy (VGG19) is 0.00% - Dog Accuracy (VGG19) is 100.00%

```
In [10]: ### TODO: Test the performance of the dog_detector function
         ### on the images in human_files_short and dog_files_short.
         vgg_human_acc = 0.0
         vgg_dog_acc = 0.0
         for human, dog in zip(human_files_short, dog_files_short):
             if(dog_detector(human)):
                 vgg_human_acc += 1
             if(dog_detector(dog)):
                 vgg_dog_acc += 1
         else:
             print('Human Accuracy (VGG19) is {:.2f}% \t Dog Accuracy (VGG19) is {:.2f}%'
                 .format(
                     vgg_human_acc/len(human_files_short) * 100,
                     vgg_dog_acc/len(dog_files_short) * 100
                  )
             )
```

```
Human Accuracy (VGG19) is 0.00%          Dog Accuracy (VGG19) is 100.00%
```

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 [11]: ### (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 [12]:

```
import os
from torchvision import datasets

### TODO: Write data loaders for training, validation, and test sets
## Specify appropriate transforms, and batch_sizes
data_dir = '/data/dog_images'
batch_size = 20
num_workers= 0
train_transforms = transforms.Compose([
    transforms.Resize(size=224),
    transforms.CenterCrop((224,224)),
    transforms.RandomHorizontalFlip(),
    transforms.RandomVerticalFlip(),
    transforms.RandomRotation(30),
    transforms.ToTensor(),
    transforms.Normalize([0.485, 0.456, 0.406],
                         [0.229, 0.224, 0.225])
])

test_transforms = transforms.Compose([
    transforms.Resize(size=224),
    transforms.CenterCrop((224,224)),
    transforms.ToTensor(),
```

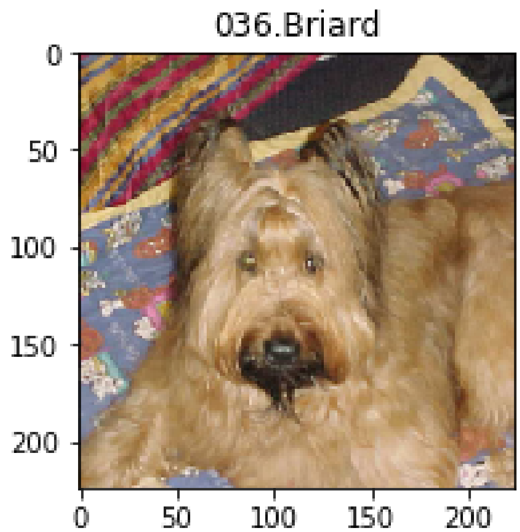
```
        transforms.Normalize([0.485, 0.456, 0.406],
                             [0.229, 0.224, 0.225])
    ])

    train_data = datasets.ImageFolder(data_dir + '/train', transform=train_transforms)
    test_data = datasets.ImageFolder(data_dir + '/test', transform=test_transforms)
    valid_data = datasets.ImageFolder(data_dir + '/valid', transform=test_transforms)

    print('  Number of train images: ', len(train_data))
    print('  Number of test images:  ', len(test_data))
    print('  Number of valid images: ', len(valid_data))

    train_loader = torch.utils.data.DataLoader(train_data, batch_size=batch_size, num_worke
    test_loader = torch.utils.data.DataLoader(test_data, batch_size=batch_size, num_workers
    valid_loader = torch.utils.data.DataLoader(valid_data, batch_size=batch_size, num_worke

    loaders_scratch = {}
    loaders_scratch['train'] = train_loader
    loaders_scratch['test'] = test_loader
    loaders_scratch['valid'] = valid_loader
```

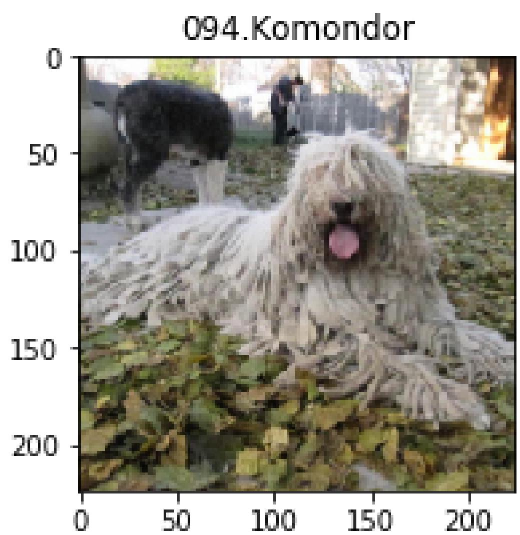
```
  Number of train images:  6680
  Number of test images:   836
  Number of valid images:  835
```

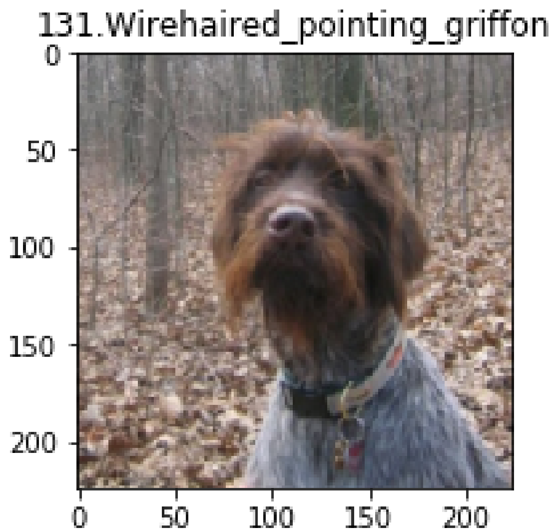
In [13]:
```
class_names = train_data.classes

print("number of classes:", len(class_names))
```

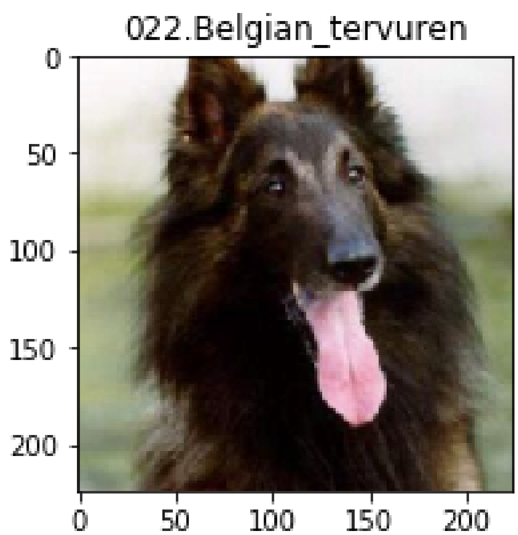
```
number of classes: 133
```

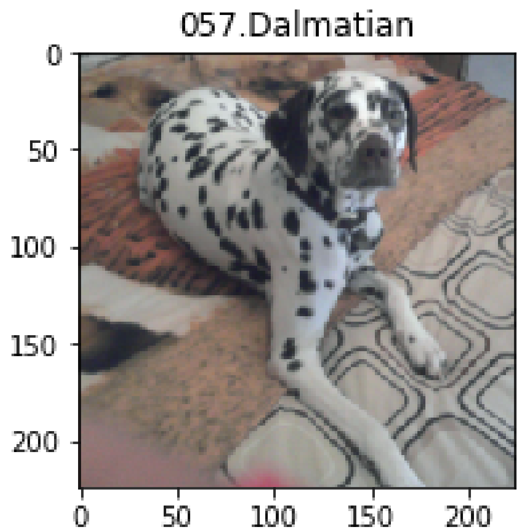
In [14]:
```
# display test data
inputs, classes = next( iter(loaders_scratch['test']) )

for image, label in zip(inputs, classes):
    image = image.to("cpu").clone().detach()
    image = image.numpy().squeeze()
    image = image.transpose(1,2,0)
    # normalize image
    image = image * np.array((0.229, 0.224, 0.225)) + np.array((0.485, 0.456, 0.406))
    image = image.clip(0, 1)

    fig = plt.figure(figsize=(12,3))
    plt.imshow(image)
    plt.title(class_names[label.item()])
```

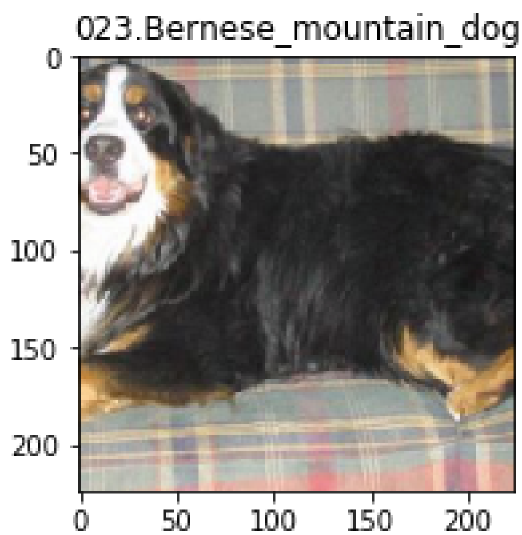

036.Briard
0
50
100
150
200
0
50
100
150
200

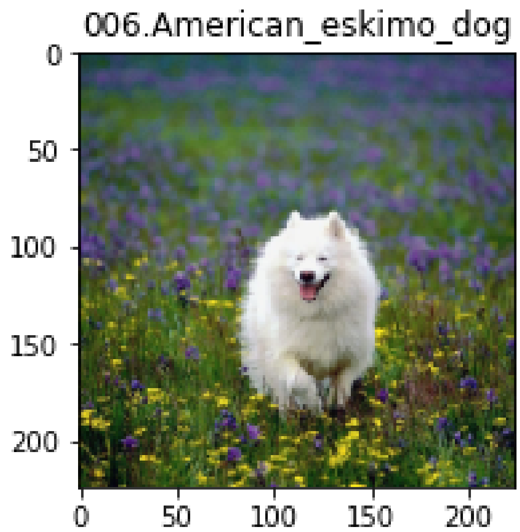

094.Komondor
0
50
100
150
200
0
50
100
150
200

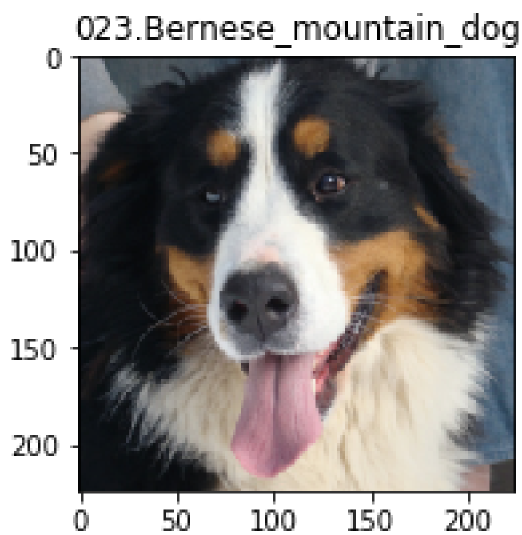

131.Wirehaired_pointing_griffon
0
50
100
150
200
0 50 100 150 200

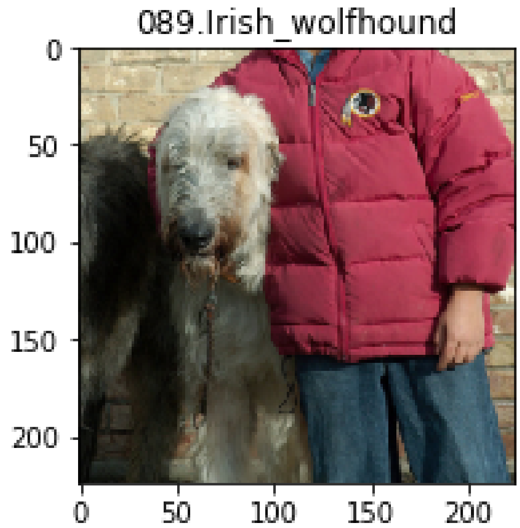

022.Belgian_tervuren
0
50
100
150
200
0 50 100 150 200

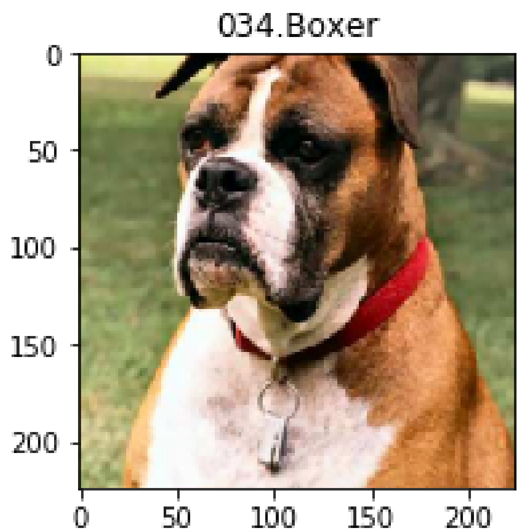

057.Dalmatian
0
50
100
150
200
0
50
100
150
200

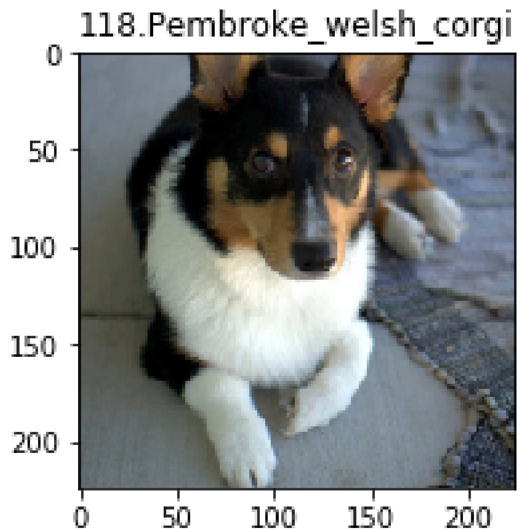

023.Bernese_mountain_dog
0
50
100
150
200
0
50
100
150
200

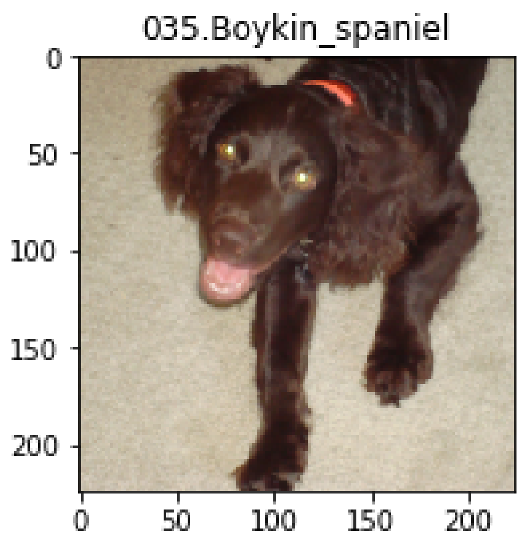

006.American_eskimo_dog
0
50
100
150
200
0
50
100
150
200

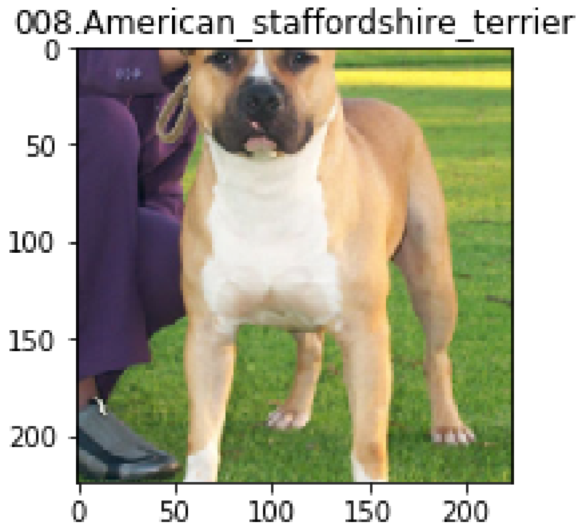

023.Bernese_mountain_dog
0
50
100
150
200
0
50
100
150
200

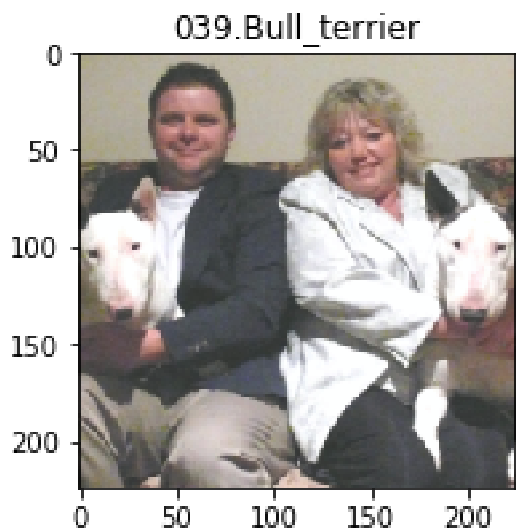

089.Irish_wolfhound
0
50
100
150
200
0
50
100
150
200

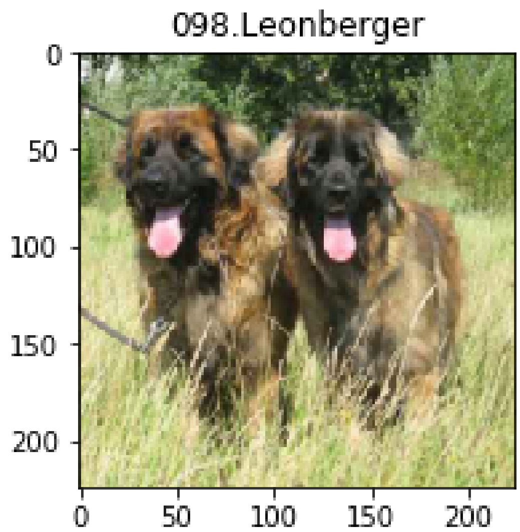

034.Boxer
0
50
100
150
200
0
50
100
150
200

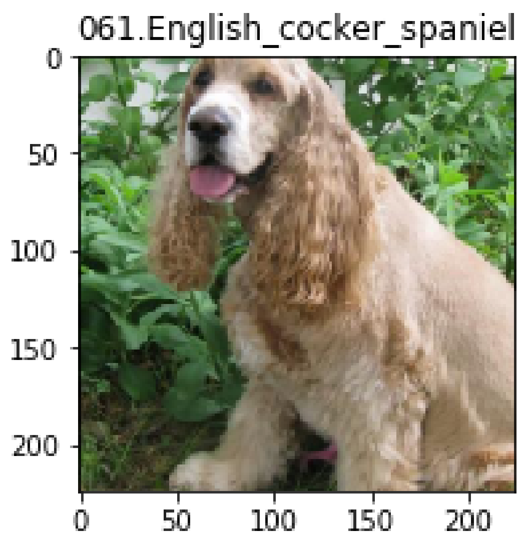

118.Pembroke_welsh_corgi
0
50
100
150
200
0
50
100
150
200

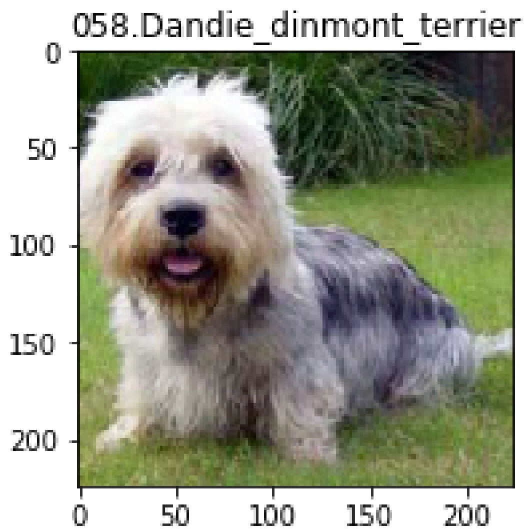

035.Boykin_spaniel
0
50
100
150
200
0
50
100
150
200

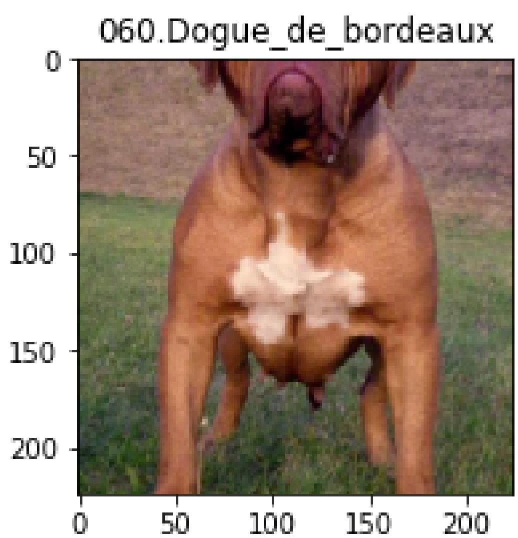

008.American_staffordshire_terrier
0
50
100
150
200

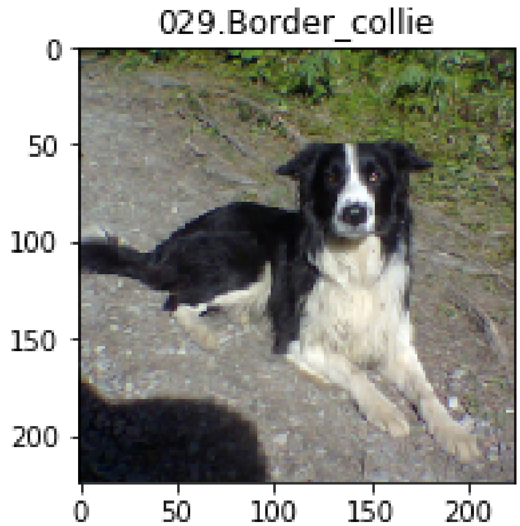

039.Bull_terrier
0
50
100
150
200

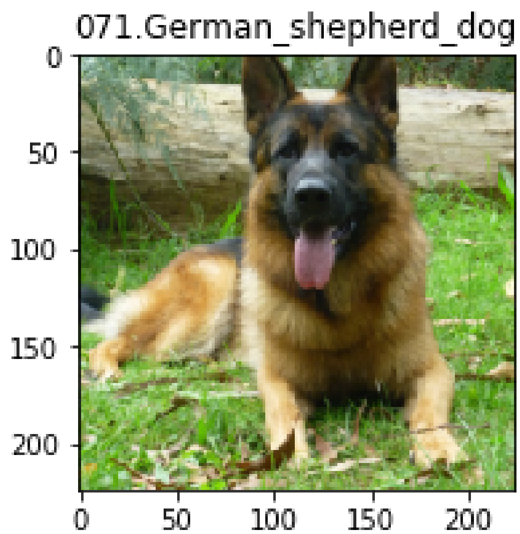

098.Leonberger
0
50
100
150
200
0 50 100 150 200


061.English_cocker_spaniel
0
50
100
150
200
0 50 100 150 200


058.Dandie_dinmont_terrier
0
50
100
150
200
0
50
100
150
200


060.Dogue_de_bordeaux
0
50
100
150
200
0
50
100
150
200





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 resize the images into a 224x224 pixels and i perform a center crop of 224 on both train, test and valid datasets - I also augmented the train data set with Random Horizontal Flip, Random Vertical Flip, Random flips of 30 degress in any direction.

1.1.8 (IMPLEMENTATION) Model Architecture

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

In [15]:
```
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 224x224x3 image tensor)
        self.conv1 = nn.Conv2d(3, 16, 3, padding=1)
        # convolutional layer (sees 112x112x16 image tensor)
        self.conv2 = nn.Conv2d(16, 32, 3, padding=1)
        # convolutional layer (sees 56x56x32 image tensor)
        self.conv3 = nn.Conv2d(32, 64, 3, padding=1)
        # convolutional layer (sees 28x28x64 image tensor)
        self.conv4 = nn.Conv2d(64, 128, 3, padding=1)

        self.maxpool = nn.MaxPool2d(2, 2)

        self.fc1 = nn.Linear(14 * 14 * 128, 1024)
        self.fc2 = nn.Linear(1024, 512)
        self.fc3 = nn.Linear(512, 256)
        self.fc4 = nn.Linear(256, 133)

        self.dropout = nn.Dropout(0.25)
        self.batch_norm = nn.BatchNorm1d(num_features=1024)

    def forward(self, x):
        ## Define forward behavior
        x = self.maxpool(F.relu(self.conv1(x)))
        x = self.maxpool(F.relu(self.conv2(x)))
        x = self.maxpool(F.relu(self.conv3(x)))
        x = self.maxpool(F.relu(self.conv4(x)))

        x = x.view(x.size(0), -1)

        x = F.relu(self.batch_norm(self.fc1(x)))
        x = self.dropout(x)
        x = F.relu(self.fc2(x))
        x = self.dropout(x)
        x = F.relu(self.fc3(x))
        x = self.dropout(x)
        x = self.fc4(x)
        return x

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

```
    # instantiate the CNN
    model_scratch = Net()
    print(model_scratch)

    # move tensors to GPU if CUDA is available
    if use_cuda:
        model_scratch.cuda()
```

```
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))
  (conv4): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (maxpool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (fc1): Linear(in_features=25088, out_features=1024, bias=True)
  (fc2): Linear(in_features=1024, out_features=512, bias=True)
  (fc3): Linear(in_features=512, out_features=256, bias=True)
  (fc4): Linear(in_features=256, out_features=133, bias=True)
  (dropout): Dropout(p=0.25)
  (batch_norm): BatchNorm1d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True
)
```

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

Answer: - The NN feature extraction layer includes a 4 layers Convolutional layers each followed by a Relu Activation function and a MaxPooling layer. - The classification aspect of the NN includes 3 fully connected layers each followed by a Relu Activation funcion and a dropout of 25%

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 [16]: 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)
         print(optimizer_scratch)
```

```
SGD (
Parameter Group 0
    dampening: 0
    lr: 0.01
    momentum: 0
```

```
    nesterov: False
    weight_decay: 0
)
```

```
In [17]: #important for trauncated images
         from PIL import ImageFile
         ImageFile.LOAD_TRUNCATED_IMAGES = True
```

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 [18]: 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

             if os.path.exists(save_path):
                     print("load previous saved model ...")
                     model.load_state_dict(torch.load(save_path))

             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()
                     # clear the gradients
                     optimizer.zero_grad()
                     # forward pass
                     output = model(data)
                     # calculate loss
                     loss = criterion(output, target)
                     # backward pass
                     loss.backward()
                     # perfrom step operation to update values
                     optimizer.step()
                     ## 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
                train_loss += loss.item() * data.size(0)

            ######################
            # 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)
                # calculate loss
                loss = criterion(output, target)

                valid_loss += loss.item() * data.size(0)

            # calculate average losses
            train_loss = train_loss / len(loaders['train'].dataset)
            valid_loss = valid_loss / len(loaders['valid'].dataset)


            # print training/validation statistics
            print('Epoch: {} \tTraining Loss: {:.6f} \tValidation Loss: {:.6f}'.format(
                epoch,
                train_loss,
                valid_loss
                ), end="")

            ## TODO: save the model if validation loss has decreased
            if(valid_loss <= valid_loss_min):
                print('  Saving model ...')
                valid_loss_min = valid_loss
                torch.save(model.state_dict(), save_path)
            else:
                print("")
        # return trained model
        return model
```

```
In [19]: # train the model
         model_scratch = train(5, loaders_scratch, model_scratch, optimizer_scratch,
                               criterion_scratch, use_cuda, 'model_scratch.pt')
```

```
load previous saved model ...
Epoch: 1         Training Loss: 3.381198         Validation Loss: 3.622691  Saving model ...
Epoch: 2         Training Loss: 3.353159         Validation Loss: 3.686537
Epoch: 3         Training Loss: 3.317421         Validation Loss: 3.464659  Saving model ...
```

```
Epoch: 4        Training Loss: 3.260372        Validation Loss: 3.432937  Saving model ...
Epoch: 5        Training Loss: 3.227897        Validation Loss: 3.424049  Saving model ...
```

```
In [20]: # 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 [21]: 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 [22]: # call test function
         test(loaders_scratch, model_scratch, criterion_scratch, use_cuda)
```

```
Test Loss: 3.459080


Test Accuracy: 16% (141/836)
```

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 [23]: ## TODO: Specify data loaders
         loaders_transfer = loaders_scratch
         loaders_transfer
```

```
Out[23]: {'train': <torch.utils.data.dataloader.DataLoader at 0x7fa34928eba8>,
          'test': <torch.utils.data.dataloader.DataLoader at 0x7fa34928ecc0>,
          'valid': <torch.utils.data.dataloader.DataLoader at 0x7fa34928ee48>}
```

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

         ## TODO: Specify model architecture
         model_transfer = models.vgg19(pretrained=True)

         ## freeze weights in features layers only
         for param in model_transfer.features.parameters():
             param.requires_grad = False

         ## change the last FCL in model
         model_transfer.classifier[6] = nn.Linear(model_transfer.classifier[6].in_features, 133)
         print(model_transfer)


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

```
Downloading: "https://download.pytorch.org/models/vgg19-dcbb9e9d.pth" to /root/.torch/models/vgg
100%|| 574673361/574673361 [00:05<00:00, 104364413.25it/s]
```

```
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): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (17): ReLU(inplace)
    (18): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    (19): Conv2d(256, 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): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (24): ReLU(inplace)
    (25): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (26): ReLU(inplace)
    (27): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    (28): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (29): ReLU(inplace)
    (30): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (31): ReLU(inplace)
    (32): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (33): ReLU(inplace)
    (34): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (35): ReLU(inplace)
    (36): 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=133, 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 downloaded the vgg19 model from the pytorch.models module. I can see the vgg19 has 2 layers namely the features and classifier layers. So i decided to freeze the weights in the features layer and use it has a feature extractor. In the classifier layer, there is a series of Linear layers and i changed the last Linear layer to one that has my proposed outcome instead of `1000` outcomes. I would be fine-tuing the weights in the classifier layer.

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 [25]: criterion_transfer = nn.CrossEntropyLoss()
         optimizer_transfer = optim.SGD(model_transfer.classifier.parameters(), lr=0.01)
```

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 [26]: # train the model
         model_transfer = train(5, loaders_transfer, model_transfer, optimizer_transfer, criteri
```

```
load previous saved model ...
Epoch: 1        Training Loss: 0.825841        Validation Loss: 0.443726  Saving model ...
Epoch: 2        Training Loss: 0.778977        Validation Loss: 0.492761
Epoch: 3        Training Loss: 0.727069        Validation Loss: 0.496448
Epoch: 4        Training Loss: 0.709552        Validation Loss: 0.484374
Epoch: 5        Training Loss: 0.657341        Validation Loss: 0.474488
```

```
In [27]: # 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 [28]: test(loaders_transfer, model_transfer, criterion_transfer, use_cuda)
```

```
Test Loss: 0.522444


Test Accuracy: 84% (709/836)
```

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 [29]: ### 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 train_data.classes]

         def predict_breed_transfer(img_path):
             image_tensor = image_to_tensor(img_path)
             if use_cuda:
                 image_tensor = image_tensor.cuda()
             output = model_transfer(image_tensor)
             _, prob_c = output.topk(1, dim=1)
             prob_c = np.squeeze(prob_c.numpy()) if not use_cuda else np.squeeze(prob_c.cpu().nu
             # load the image and return the predicted breed
             return class_names[prob_c]

         def display_image(img_path, title="Title"):
             image = Image.open(img_path)
             plt.title(title)
             plt.imshow(image)
             plt.show()
```

Step 5: Write your Algorithm

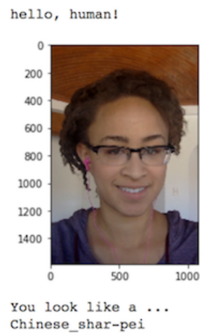
Write an algorithm that accepts a file path to an image and first determines whether the image contains a human, dog, or neither. Then, - if a **dog** is detected in the image, return the predicted breed. - if a **human** is detected in the image, return the resembling dog breed. - if **neither** is 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 [30]: ### 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
             image_tensor = image_to_tensor(img_path)
```



Sample Human Output

```
if(face_detector(img_path)):
    print("Human Face Dected!")
    predicted_breed = predict_breed_transfer(img_path)
    display_image(img_path, title="Predicted: {}".format(predicted_breed) )
    print("You look like a ...")
    print(predicted_breed)
elif(dog_detector(img_path)):
    print("Dog Face Detected!")
    predicted_breed = predict_breed_transfer(img_path)
    display_image(img_path, title="Predicted: {}".format(predicted_breed) )
    print("Its a ...")
    print(predicted_breed)

else:
    print("Oh, Sorry! Couldn't detect any dog or human face in the image.")
    display_image(img_path, title="...")
    print("Try another!")

print("\n")
```

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.

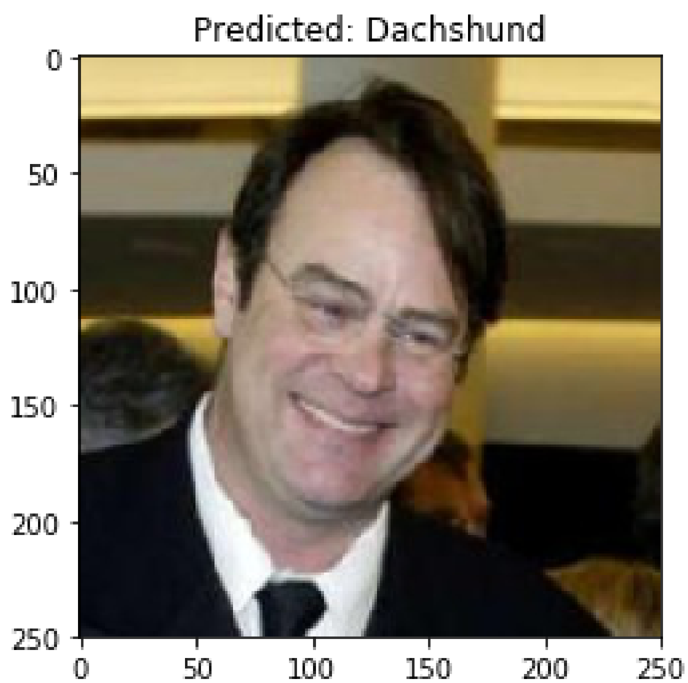
Question 6: Is the output better than you expected :) ? Or worse :(? Provide at least three possible points of improvement for your algorithm.

Answer:

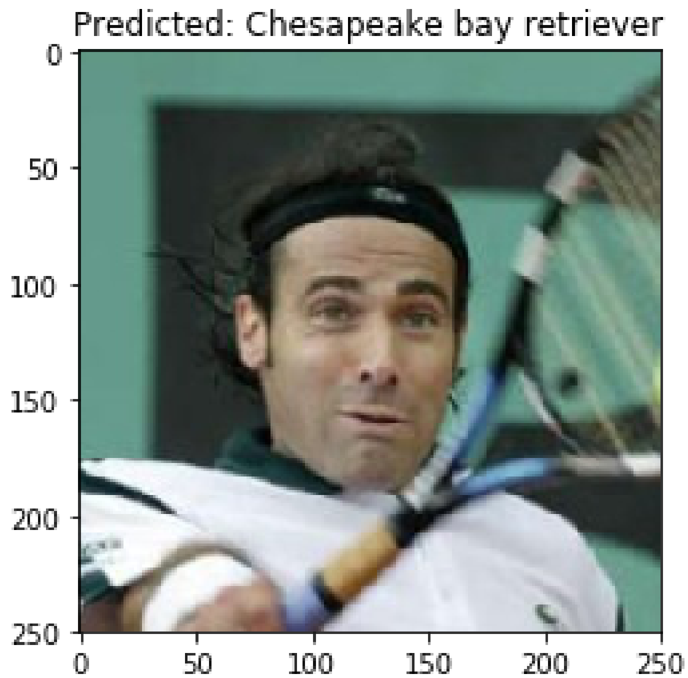
- Fine tune the model more and also increase training time.
- Implement and provide this model as an API using flask.
- Revamp and clean up the code
- Implement with different training parameters (lr, optimizer etc)

```
In [31]: ## 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)
```

```
Human Face Dected!
```

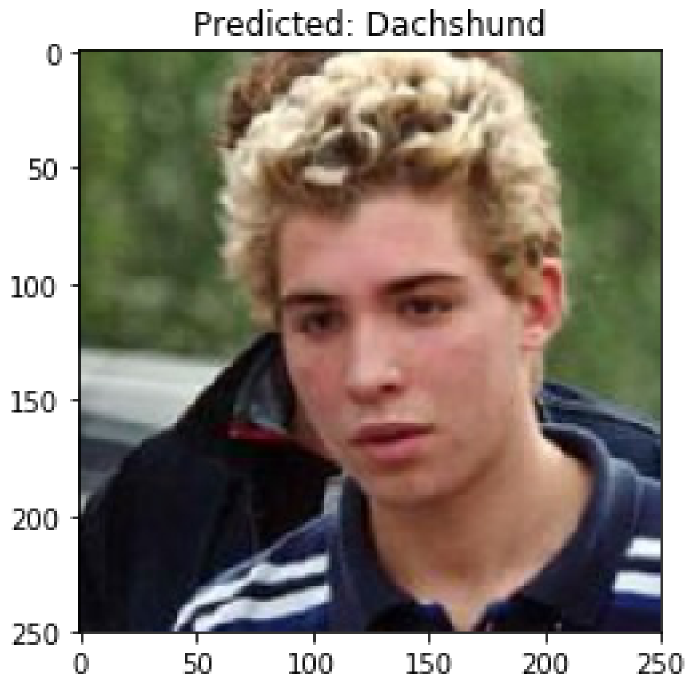



```
You look like a ...
Dachshund


Human Face Dected!
```



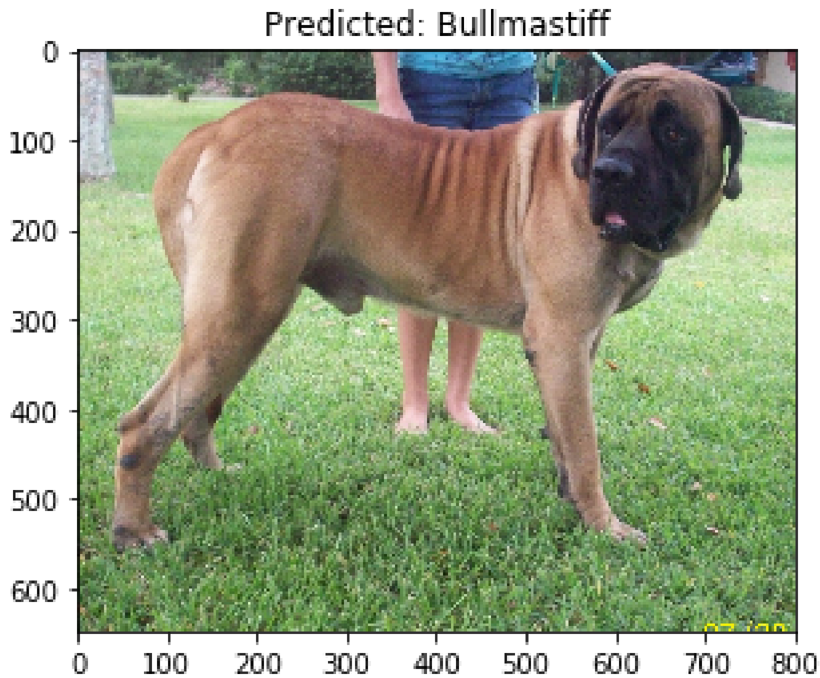

```
You look like a ...
Chesapeake bay retriever
```

```
Human Face Dected!
```



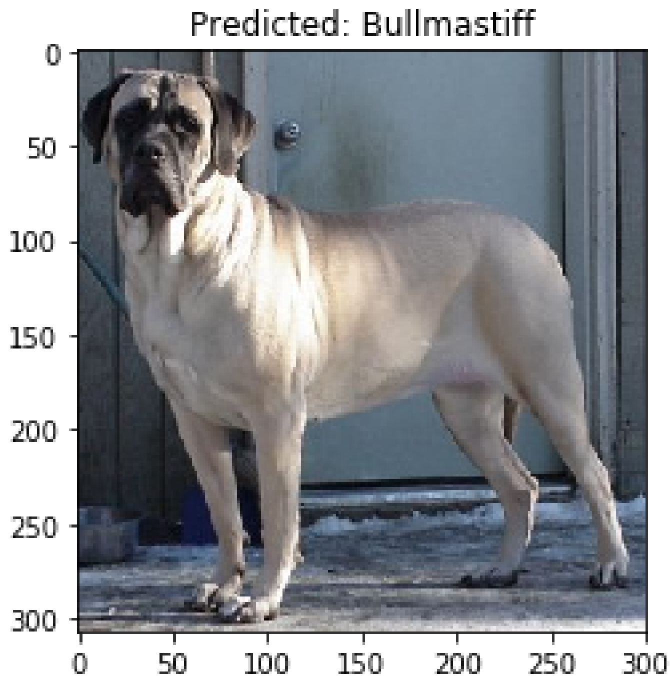

```
You look like a ...
Dachshund
```

```
Dog Face Detected!
```

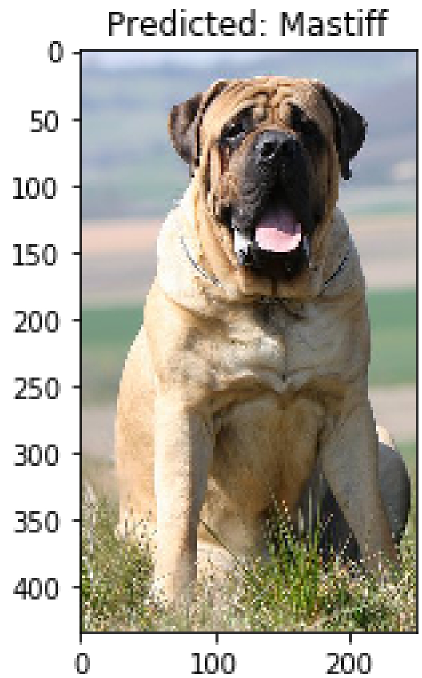



```
Its a ...
Bullmastiff
```

```
Dog Face Detected!
```



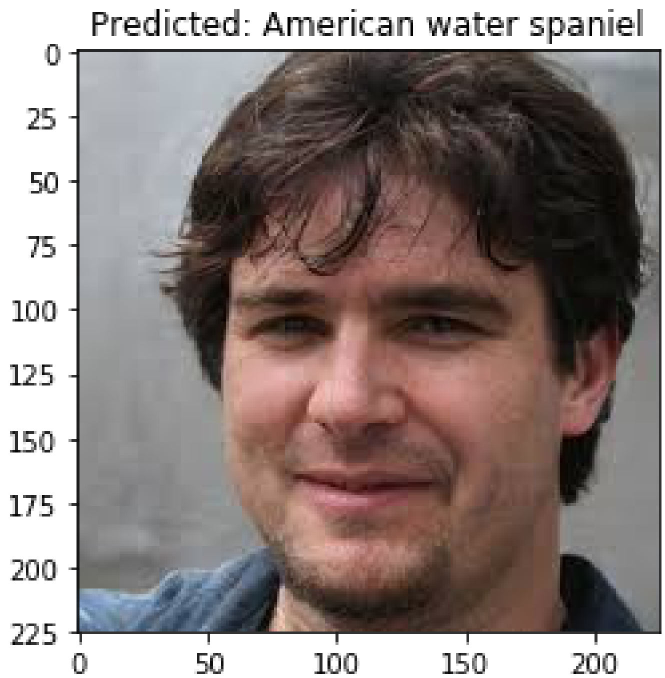

```
Its a ...
Bullmastiff


Dog Face Detected!
```




```
Its a ...
Mastiff
```

```
In [32]: # load local images
         local_test_data = np.array(glob("./local_test_data/*"))
         for file in local_test_data:
             run_app(file)
```

```
Human Face Dected!
```

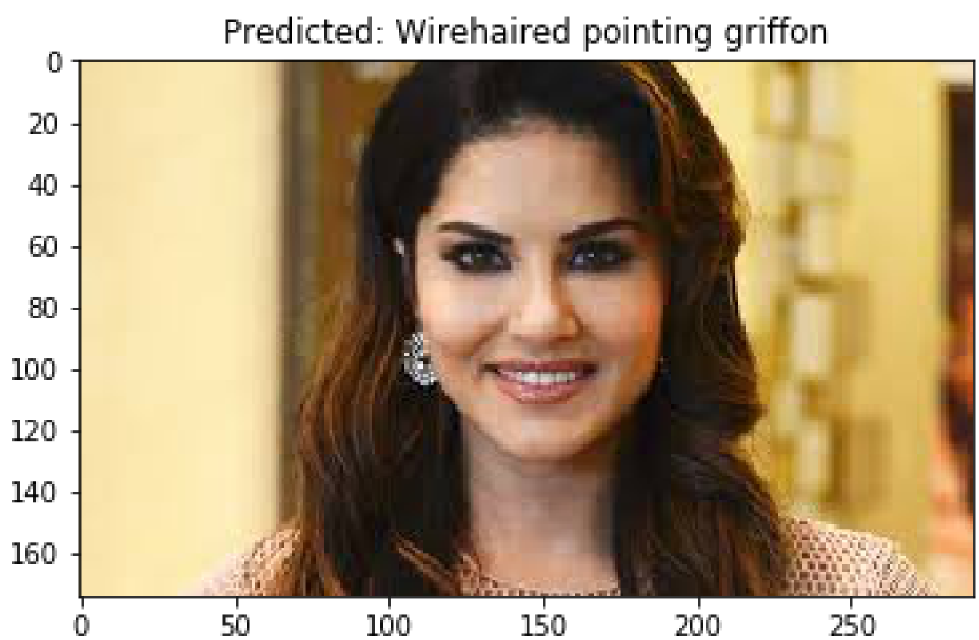



```
You look like a ...
American water spaniel


Human Face Dected!
```


```
You look like a ...
Dandie dinmont terrier
```

```
Human Face Dected!
```


```
You look like a ...
Belgian sheepdog


Human Face Dected!
```



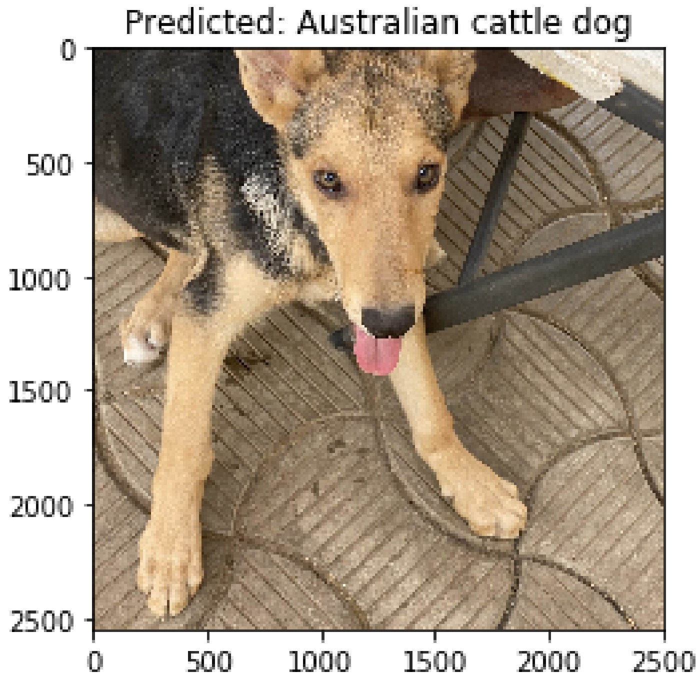
```
You look like a ...
Wirehaired pointing griffon
```

```
Human Face Dected!
```


```
You look like a ...
Dogue de bordeaux
```

```
Dog Face Detected!
```


```
Its a ...
Labrador retriever


Dog Face Detected!
```




```
Its a ...
Australian cattle dog
```

```
In [ ]:
```