Artificial Intelligence Nanodegree

Convolutional Neural Networks

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 iPython 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 \n", "**File -> Download as -> HTML (.html)**. Include the finished document along with this notebook as your submission.

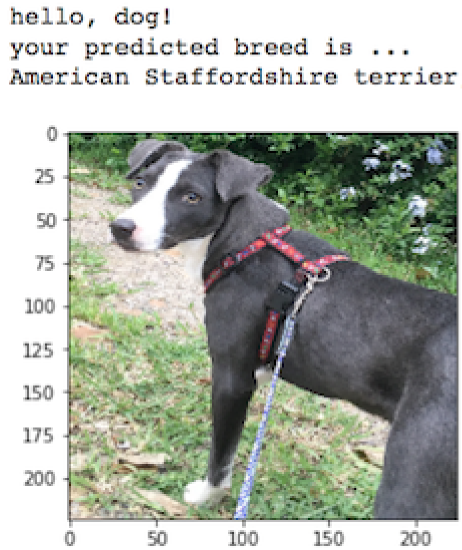
In addition to implementing code, there will be questions that you must answer which relate to the project and your implementation. Each section where you will answer a question is preceded by a **'Question X'** header. Carefully read each question and provide thorough answers in the following text boxes 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 IPython notebook.

Why We're Here

In this notebook, you will make the first steps towards developing an algorithm that could be used as part of a mobile or web app. At the end of this project, your code will accept any user-supplied image as input. If a dog is detected in the image, it will provide an estimate of the dog's breed. If a human is detected, it will provide an estimate of the dog breed that is most resembling. The image below displays potential sample output of your finished project (... but we expect that each student's algorithm will behave differently!).



In this real-world setting, you will need to piece together a series of models to perform different tasks; for instance, the algorithm that detects humans in an image will be different from the CNN that infers dog breed. There are many points of possible failure, and no perfect algorithm exists. Your imperfect solution will nonetheless create a fun user experience!

The Road Ahead

We break the notebook into separate steps. Feel free to use the links below to navigate the notebook.

- Step 0: Import Datasets
- Step 1: Detect Humans
- Step 2: Detect Dogs
- Step 3: Create a CNN to Classify Dog Breeds (from Scratch)
- Step 4: Use a CNN to Classify Dog Breeds (using Transfer Learning)
- Step 5: Create a CNN to Classify Dog Breeds (using Transfer Learning)
- Step 6: Write your Algorithm
- Step 7: Test Your Algorithm

Step 0: Import Datasets

Import Dog Dataset

In the code cell below, we import a dataset of dog images. We populate a few variables through the use of the `load_files` function from the scikit-learn library:

- `train_files`, `valid_files`, `test_files` - numpy arrays containing file paths to images
- `train_targets`, `valid_targets`, `test_targets` - numpy arrays containing onehot-encoded classification labels
- `dog_names` - list of string-valued dog breed names for translating labels

In [1]:
```
from sklearn.datasets import load_files
from keras.utils import np_utils
import numpy as np
from glob import glob

# define function to load train, test, and validation datasets
def load_dataset(path):
    data = load_files(path)
    dog_files = np.array(data['filenames'])
    dog_targets = np_utils.to_categorical(np.array(data['target']), 1
33)
    return dog_files, dog_targets

# load train, test, and validation datasets
train_files, train_targets = load_dataset('dogImages/train')
valid_files, valid_targets = load_dataset('dogImages/valid')
test_files, test_targets = load_dataset('dogImages/test')

# load list of dog names
dog_names = [item[20:-1] for item in
sorted(glob("dogImages/train/*/"))]

# print statistics about the dataset
print('There are %d total dog categories.' % len(dog_names))
print('There are %s total dog images.\n' %
len(np.hstack([train_files, valid_files, test_files])))
print('There are %d training dog images.' % len(train_files))
print('There are %d validation dog images.' % len(valid_files))
print('There are %d test dog images.'% len(test_files))
```

```
Using TensorFlow backend.

There are 133 total dog categories.
There are 8351 total dog images.

There are 6680 training dog images.
There are 835 validation dog images.
There are 836 test dog images.
```

Import Human Dataset

In the code cell below, we import a dataset of human images, where the file paths are stored in the numpy array `human_files`.

In [2]:
```
import random
random.seed(8675309)

# load filenames in shuffled human dataset
human_files = np.array(glob("lfw/*/*"))
random.shuffle(human_files)

# print statistics about the dataset
print('There are %d total human images.' % len(human_files))
```

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

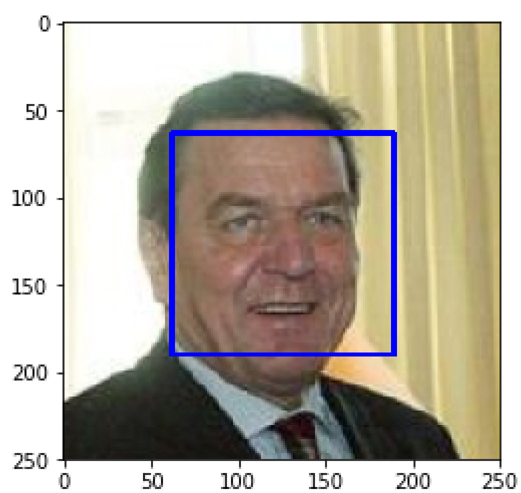
Step 1: Detect Humans

We use OpenCV's implementation of Haar feature-based cascade classifiers (http://docs.opencv.org/trunk/d7/d8b/tutorial_py_face_detection.html) to detect human faces in images. OpenCV provides many pre-trained face detectors, stored as XML files on github (https://github.com/opencv/opencv/tree/master/data/haarcascades). 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 [3]:
```
import cv2
import matplotlib.pyplot as plt
%matplotlib inline

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

# load color (BGR) image
img = cv2.imread(human_files[3])
# 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.

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 [4]:
```
# 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
```

(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:

In [5]:
```
human_files_short = human_files[:100]
dog_files_short = train_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.

face_human_detect = 0
face_dog_detect = 0

for _,face in enumerate(human_files_short):
    if face_detector(face):
        face_human_detect += 1

for _,face in enumerate(dog_files_short):
    if face_detector(face):
        face_dog_detect += 1

print("Using Human face detector...")
print("Percentage of human faces identified:",face_human_detect/len(h
uman_files_short)*100)
print("Percentage of dogs identified as
humans:",face_dog_detect/len(dog_files_short)*100)
```

```
Using Human face detector...
Percentage of human faces identified: 100.0
Percentage of dogs identified as humans: 11.0
```

Question 2: This algorithmic choice necessitates that we communicate to the user that we accept human images only when they provide a clear view of a face (otherwise, we risk having unneccessarily frustrated users!). In your opinion, is this a reasonable expectation to pose on the user? If not, can you think of a way to detect humans in images that does not necessitate an image with a clearly presented face?

Answer: It is not practical to pose limitations to the face images that the user should provide. Apart from the inconvenience to the user, when the result is negative the user will not know whether it is actually negative or a false negative.

Apart from the face there are other characteristics of humans that can be detected, such as the body characteristics. Using only face dtection algorithm to detect humans can capture only a part of possible outcomes.

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 each of the datasets.

In [6]:
```
## (Optional) TODO: Report the performance of another
## face detection algorithm on the LFW dataset
### Feel free to use as many code cells as needed.
```

Step 2: Detect Dogs

In this section, we use a pre-trained [ResNet-50 (http://ethereon.github.io/netscope/#/gist/db945b393d40bfa26006)] model to detect dogs in images. Our first line of code downloads the ResNet-50 model, along with weights that have been trained on ImageNet (http://www.image-net.org/), 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 (https://gist.github.com/yrevar/942d3a0ac09ec9e5eb3a). Given an image, this pre-trained ResNet-50 model returns a prediction (derived from the available categories in ImageNet) for the object that is contained in the image.

In [7]:
```
from keras.applications.resnet50 import ResNet50

# define ResNet50 model
ResNet50_model = ResNet50(weights='imagenet')
```

Pre-process the Data

When using TensorFlow as backend, Keras CNNs require a 4D array (which we'll also refer to as a 4D tensor) as input, with shape

$$(\text{nb\_samples}, \text{rows}, \text{columns}, \text{channels}),$$

where `nb_samples` corresponds to the total number of images (or samples), and `rows`, `columns`, and `channels` correspond to the number of rows, columns, and channels for each image, respectively.

The `path_to_tensor` function below takes a string-valued file path to a color image as input and returns a 4D tensor suitable for supplying to a Keras CNN. The function first loads the image and resizes it to a square image that is 224×224 pixels. Next, the image is converted to an array, which is then resized to a 4D tensor. In this case, since we are working with color images, each image has three channels. Likewise, since we are processing a single image (or sample), the returned tensor will always have shape

$$(1, 224, 224, 3).$$

The `paths_to_tensor` function takes a numpy array of string-valued image paths as input and returns a 4D tensor with shape

$$(\text{nb\_samples}, 224, 224, 3).$$

Here, `nb_samples` is the number of samples, or number of images, in the supplied array of image paths. It is best to think of `nb_samples` as the number of 3D tensors (where each 3D tensor corresponds to a different image) in your dataset!

In [8]:
```
from keras.preprocessing import image
from tqdm import tqdm

def path_to_tensor(img_path):
    # loads RGB image as PIL.Image.Image type
    img = image.load_img(img_path, target_size=(224, 224))
    # convert PIL.Image.Image type to 3D tensor with shape (224, 224, 3)
    x = image.img_to_array(img)
    # convert 3D tensor to 4D tensor with shape (1, 224, 224, 3) and return 4D tensor
    return np.expand_dims(x, axis=0)

def paths_to_tensor(img_paths):
    list_of_tensors = [path_to_tensor(img_path) for img_path in tqdm(img_paths)]
    return np.vstack(list_of_tensors)
```

Making Predictions with ResNet-50

Getting the 4D tensor ready for ResNet-50, and for any other pre-trained model in Keras, requires some additional processing. First, the RGB image is converted to BGR by reordering the channels. All pre-trained models have the additional normalization step that the mean pixel (expressed in RGB as $[103.939, 116.779, 123.68]$ and calculated from all pixels in all images in ImageNet) must be subtracted from every pixel in each image. This is implemented in the imported function `preprocess_input`. If you're curious, you can check the code for `preprocess_input` here (https://github.com/fchollet/keras/blob/master/keras/applications/imagenet_utils.py).

Now that we have a way to format our image for supplying to ResNet-50, we are now ready to use the model to extract the predictions. This is accomplished with the `predict` method, which returns an array whose i-th entry is the model's predicted probability that the image belongs to the i-th ImageNet category. This is implemented in the `ResNet50_predict_labels` function below.

By taking the argmax of the predicted probability vector, we obtain an integer corresponding to the model's predicted object class, which we can identify with an object category through the use of this dictionary (https://gist.github.com/yrevar/942d3a0ac09ec9e5eb3a).

In [9]:
```
from keras.applications.resnet50 import preprocess_input, decode_predictions

def ResNet50_predict_labels(img_path):
    # returns prediction vector for image located at img_path
    img = preprocess_input(path_to_tensor(img_path))
    return np.argmax(ResNet50_model.predict(img))
```

Write a Dog Detector

While looking at the dictionary (https://gist.github.com/yrevar/942d3a0ac09ec9e5eb3a), 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 ResNet-50 model, we need only check if the `ResNet50_predict_labels` function above returns a value between 151 and 268 (inclusive).

We 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 [10]:
```
### returns "True" if a dog is detected in the image stored at img_pa
th
def dog_detector(img_path):
    prediction = ResNet50_predict_labels(img_path)
    return ((prediction <= 268) & (prediction >= 151))
```

(IMPLEMENTATION) Assess the Dog Detector

Question 3: Use the code cell below to test the performance of your `dog_detector` function.

- What percentage of the images in `human_files_short` have a detected dog?
- What percentage of the images in `dog_files_short` have a detected dog?

Answer:

In [11]:

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

face_human_detect = 0
face_dog_detect = 0

# check for human faces with dog detector
for _,face in enumerate(human_files_short):
    if dog_detector(face):
        face_human_detect += 1

# check for dogs with dog detector
for _,face in enumerate(dog_files_short):
    if dog_detector(face):
        face_dog_detect += 1

print("Using dog detector...")
print("Percentage of human faces identified as dogs:",face_human_dete
ct/len(human_files_short)*100)
print("Percentage of dogs identifies:",face_dog_detect/len(dog_files_
short)*100)
```

```
Using dog detector...
Percentage of human faces identified as dogs: 1.0
Percentage of dogs identifies: 100.0
```

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

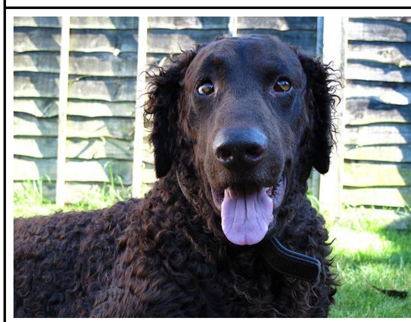
Now that we have functions for detecting humans and dogs in images, we need a way to predict breed from images. In this step, you will create a CNN that classifies dog breeds. You must create your CNN *from scratch* (so, you can't use transfer learning *yet*!), and you must attain a test accuracy of at least 1%. In Step 5 of this notebook, you will have the opportunity to use transfer learning to create a CNN that attains greatly improved accuracy.

Be careful with adding too many trainable layers! More parameters means longer training, which means you are more likely to need a GPU to accelerate the training process. Thankfully, Keras provides a handy estimate of the time that each epoch is likely to take; you can extrapolate this estimate to figure out how long it will take for your algorithm to train.

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 great difficulty in 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**	**Black 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!

Pre-process the Data

We rescale the images by dividing every pixel in every image by 255.

In [12]:
```
from PIL import ImageFile
ImageFile.LOAD_TRUNCATED_IMAGES = True

# pre-process the data for Keras
train_tensors = paths_to_tensor(train_files).astype('float32')/255
valid_tensors = paths_to_tensor(valid_files).astype('float32')/255
test_tensors = paths_to_tensor(test_files).astype('float32')/255
```

```
100%|██████████| 6680/6680 [00:42<00:00, 158.36it/s]
100%|██████████| 835/835 [00:04<00:00, 175.18it/s]
100%|██████████| 836/836 [00:04<00:00, 177.75it/s]
```

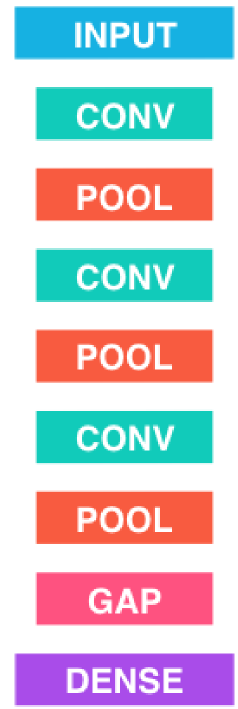
(IMPLEMENTATION) Model Architecture

Create a CNN to classify dog breed. At the end of your code cell block, summarize the layers of your model by executing the line:

```
model.summary()
```

We have imported some Python modules to get you started, but feel free to import as many modules as you need. If you end up getting stuck, here's a hint that specifies a model that trains relatively fast on CPU and attains >1% test accuracy in 5 epochs:

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 223, 223, 16)      208
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 111, 111, 16)      0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 110, 110, 32)      2080
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 55, 55, 32)        0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 54, 54, 64)        8256
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 27, 27, 64)        0
_________________________________________________________________
global_average_pooling2d_1 ( (None, 64)                0
_________________________________________________________________
dense_1 (Dense)              (None, 133)               8645
=================================================================
Total params: 19,189.0
Trainable params: 19,189.0
Non-trainable params: 0.0
_________________________________________________________________
```



Question 4: Outline the steps you took to get to your final CNN architecture and your reasoning at each step. If you chose to use the hinted architecture above, describe why you think that CNN architecture should work well for the image classification task.

Answer:

The input image of size (224,224,3) is fed to a series of convolutional networks with increasing depth, 32, 64 and 128 respectively. Each convolutional layer is followed by max pool layer to avoid overfitting.

The achieved test accuracy is 9%, much higher that 1% which is the pure random choice.

In [13]:

```
from keras.layers import Conv2D, MaxPooling2D, GlobalAveragePooling2D
from keras.layers import Dropout, Flatten, Dense
from keras.models import Sequential

# model architecture
model = Sequential()
model.add(Conv2D(filters=16, kernel_size=2, padding='same', activatio
n='relu',
                        input_shape=(224, 224, 3)))
model.add(MaxPooling2D(pool_size=2))

model.add(Conv2D(filters=32, kernel_size=2, padding='same', activatio
n='relu'))
model.add(MaxPooling2D(pool_size=2))

model.add(Conv2D(filters=64, kernel_size=2, padding='same', activatio
n='relu'))
model.add(MaxPooling2D(pool_size=2))

model.add(Conv2D(filters=128, kernel_size=2, padding='same', activati
on='relu'))
model.add(MaxPooling2D(pool_size=2))

model.add(GlobalAveragePooling2D())
# Output layer
model.add(Dense(133, activation='softmax'))


model.summary()
```

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
conv2d_1 (Conv2D)            (None, 224, 224, 16)      208       
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 112, 112, 16)      0         
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 112, 112, 32)      2080      
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 56, 56, 32)        0         
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 56, 56, 64)        8256      
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 28, 28, 64)        0         
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 28, 28, 128)       32896     
_________________________________________________________________
max_pooling2d_5 (MaxPooling2 (None, 14, 14, 128)       0         
_________________________________________________________________
global_average_pooling2d_1 ( (None, 128)               0         
_________________________________________________________________
dense_1 (Dense)              (None, 133)               17157     
=================================================================
Total params: 60,597
Trainable params: 60,597
Non-trainable params: 0
_________________________________________________________________
```

Compile the Model

In [14]:
```
model.compile(optimizer='rmsprop', loss='categorical_crossentropy', m
etrics=['accuracy'])
```

(IMPLEMENTATION) Train the Model

Train your model in the code cell below. Use model checkpointing to save the model that attains the best validation loss.

You are welcome to augment the training data (https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html), but this is not a requirement.

In [15]:

```
from keras.callbacks import ModelCheckpoint

### TODO: specify the number of epochs that you would like to use to
 train the model.

epochs = 10

### Do NOT modify the code below this line.

checkpointer = ModelCheckpoint(filepath='saved_models/weights.best.fr
om_scratch.hdf5',
                               verbose=1, save_best_only=True)

model.fit(train_tensors, train_targets,
          validation_data=(valid_tensors, valid_targets),
          epochs=epochs, batch_size=20, callbacks=[checkpointer], ver
bose=1)
```

```
Train on 6680 samples, validate on 835 samples
Epoch 1/10
6640/6680 [============================>.] - ETA: 0s - loss: 4.8826 -
acc: 0.0072Epoch 00000: val_loss improved from inf to 4.86747, saving
model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.8828 - ac
c: 0.0072 - val_loss: 4.8675 - val_acc: 0.0120
Epoch 2/10
6660/6680 [============================>.] - ETA: 0s - loss: 4.8621 -
acc: 0.0114Epoch 00001: val_loss improved from 4.86747 to 4.83745, sa
ving model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.8622 - ac
c: 0.0114 - val_loss: 4.8375 - val_acc: 0.0228
Epoch 3/10
6660/6680 [============================>.] - ETA: 0s - loss: 4.8026 -
acc: 0.0185Epoch 00002: val_loss improved from 4.83745 to 4.76792, sa
ving model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.8019 - ac
c: 0.0184 - val_loss: 4.7679 - val_acc: 0.0192
Epoch 4/10
6660/6680 [============================>.] - ETA: 0s - loss: 4.7231 -
acc: 0.0284Epoch 00003: val_loss improved from 4.76792 to 4.69926, sa
ving model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.7233 - ac
c: 0.0283 - val_loss: 4.6993 - val_acc: 0.0335
Epoch 5/10
6660/6680 [============================>.] - ETA: 0s - loss: 4.6468 -
acc: 0.0338Epoch 00004: val_loss improved from 4.69926 to 4.66644, sa
ving model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.6460 - ac
c: 0.0337 - val_loss: 4.6664 - val_acc: 0.0323
Epoch 6/10
6660/6680 [============================>.] - ETA: 0s - loss: 4.5759 -
acc: 0.0426Epoch 00005: val_loss improved from 4.66644 to 4.59496, sa
ving model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.5768 - ac
c: 0.0425 - val_loss: 4.5950 - val_acc: 0.0395
Epoch 7/10
6640/6680 [============================>.] - ETA: 0s - loss: 4.5022 -
acc: 0.0435Epoch 00006: val_loss improved from 4.59496 to 4.54739, sa
ving model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.5023 - ac
c: 0.0434 - val_loss: 4.5474 - val_acc: 0.0395
Epoch 8/10
6660/6680 [============================>.] - ETA: 0s - loss: 4.4378 -
acc: 0.0529Epoch 00007: val_loss improved from 4.54739 to 4.46345, sa
ving model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.4376 - ac
c: 0.0527 - val_loss: 4.4635 - val_acc: 0.0491
Epoch 9/10
6660/6680 [============================>.] - ETA: 0s - loss: 4.3818 -
acc: 0.0524Epoch 00008: val_loss improved from 4.46345 to 4.44835, sa
ving model to saved_models/weights.best.from_scratch.hdf5
6680/6680 [==============================] - 15s - loss: 4.3814 - ac
c: 0.0524 - val_loss: 4.4484 - val_acc: 0.0467
Epoch 10/10
6660/6680 [============================>.] - ETA: 0s - loss: 4.3448 -
```

```
acc: 0.0530Epoch 00009: val_loss did not improve
6680/6680 [==============================] - 15s - loss: 4.3441 - ac
c: 0.0531 - val_loss: 4.5331 - val_acc: 0.0515
```

Out[15]:
```
<keras.callbacks.History at 0x7f1d2074bf98>
```

Load the Model with the Best Validation Loss

In [16]:
```
model.load_weights('saved_models/weights.best.from_scratch.hdf5')
```

Test the Model

Try out your model on the test dataset of dog images. Ensure that your test accuracy is greater than 1%.

In [17]:
```
# get index of predicted dog breed for each image in test set
dog_breed_predictions = [np.argmax(model.predict(np.expand_dims(tenso
r, axis=0))) for tensor in test_tensors]

# report test accuracy
test_accuracy =
100*np.sum(np.array(dog_breed_predictions)==np.argmax(test_targets, a
xis=1))/len(dog_breed_predictions)
print('Test accuracy: %.4f%%' % test_accuracy)
```

```
Test accuracy: 6.3397%
```

Step 4: Use a CNN to Classify Dog Breeds

To reduce training time without sacrificing accuracy, we show you how to train a CNN using transfer learning. In the following step, you will get a chance to use transfer learning to train your own CNN.

Obtain Bottleneck Features

In [18]:
```
bottleneck_features = np.load('bottleneck_features/DogVGG16Data.npz')
train_VGG16 = bottleneck_features['train']
valid_VGG16 = bottleneck_features['valid']
test_VGG16 = bottleneck_features['test']
```

Model Architecture

The model uses the the pre-trained VGG-16 model as a fixed feature extractor, where the last convolutional output of VGG-16 is fed as input to our model. We only add a global average pooling layer and a fully connected layer, where the latter contains one node for each dog category and is equipped with a softmax.

In [19]:
```
VGG16_model = Sequential()
VGG16_model.add(GlobalAveragePooling2D(input_shape=train_VGG16.shape[1

VGG16_model.add(Dense(133, activation='softmax'))

VGG16_model.summary()
```

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
global_average_pooling2d_2 ( (None, 512)               0         
_________________________________________________________________
dense_2 (Dense)              (None, 133)               68229     
=================================================================
Total params: 68,229
Trainable params: 68,229
Non-trainable params: 0
_________________________________________________________________
```

Compile the Model

In [20]:
```
VGG16_model.compile(loss='categorical_crossentropy', optimizer='rmspr
op', metrics=['accuracy'])
```

Train the Model

In [21]:
```
checkpointer = ModelCheckpoint(filepath='saved_models/weights.best.VG
G16.hdf5',
                               verbose=1, save_best_only=True)

VGG16_model.fit(train_VGG16, train_targets,
          validation_data=(valid_VGG16, valid_targets),
          epochs=5, batch_size=20, callbacks=[checkpointer],
verbose=1)
```

```
Train on 6680 samples, validate on 835 samples
Epoch 1/5
6380/6680 [===========================>..] - ETA: 0s - loss: 12.3216
 - acc: 0.1210Epoch 00000: val_loss improved from inf to 10.80022, sa
ving model to saved_models/weights.best.VGG16.hdf5
6680/6680 [==============================] - 1s - loss: 12.2384 - ac
c: 0.1266 - val_loss: 10.8002 - val_acc: 0.2347
Epoch 2/5
6640/6680 [============================>.] - ETA: 0s - loss: 10.4986
 - acc: 0.2714Epoch 00001: val_loss improved from 10.80022 to 10.3679
2, saving model to saved_models/weights.best.VGG16.hdf5
6680/6680 [==============================] - 1s - loss: 10.4988 - ac
c: 0.2717 - val_loss: 10.3679 - val_acc: 0.2814
Epoch 3/5
6580/6680 [============================>.] - ETA: 0s - loss: 9.9789 -
acc: 0.3266Epoch 00002: val_loss improved from 10.36792 to 9.80327, s
aving model to saved_models/weights.best.VGG16.hdf5
6680/6680 [==============================] - 1s - loss: 9.9921 - acc:
0.3262 - val_loss: 9.8033 - val_acc: 0.3329
Epoch 4/5
6440/6680 [===========================>..] - ETA: 0s - loss: 9.4646 -
acc: 0.3592Epoch 00003: val_loss improved from 9.80327 to 9.50585, sa
ving model to saved_models/weights.best.VGG16.hdf5
6680/6680 [==============================] - 1s - loss: 9.4496 - acc:
0.3605 - val_loss: 9.5059 - val_acc: 0.3425
Epoch 5/5
6520/6680 [============================>.] - ETA: 0s - loss: 9.0486 -
acc: 0.3997Epoch 00004: val_loss improved from 9.50585 to 9.21020, sa
ving model to saved_models/weights.best.VGG16.hdf5
6680/6680 [==============================] - 1s - loss: 9.0252 - acc:
0.4009 - val_loss: 9.2102 - val_acc: 0.3617
```

Out[21]:
```
<keras.callbacks.History at 0x7f1d20157780>
```

Load the Model with the Best Validation Loss

In [22]:
```
VGG16_model.load_weights('saved_models/weights.best.VGG16.hdf5')
```

Test the Model

Now, we can use the CNN to test how well it identifies breed within our test dataset of dog images. We print the test accuracy below.

In [23]:
```
# get index of predicted dog breed for each image in test set
VGG16_predictions = [np.argmax(VGG16_model.predict(np.expand_dims(fea
ture, axis=0))) for feature in test_VGG16]
# report test accuracy
test_accuracy = 100*np.sum(np.array(VGG16_predictions)==np.argmax(tes
t_targets, axis=1))/len(VGG16_predictions)
print('Test accuracy: %.4f%%' % test_accuracy)
```

```
Test accuracy: 35.1675%
```

Predict Dog Breed with the Model

In [24]:
```
from extract_bottleneck_features import *

def VGG16_predict_breed(img_path):
    # extract bottleneck features
    bottleneck_feature = extract_VGG16(path_to_tensor(img_path))
    # obtain predicted vector
    predicted_vector = VGG16_model.predict(bottleneck_feature)
    # return dog breed that is predicted by the model
    return dog_names[np.argmax(predicted_vector)]
```

Step 5: 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.

In Step 4, we used transfer learning to create a CNN using VGG-16 bottleneck features. In this section, you must use the bottleneck features from a different pre-trained model. To make things easier for you, we have pre-computed the features for all of the networks that are currently available in Keras:

- VGG-19 (https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/DogVGG19Data.npz) bottleneck features
- ResNet-50 (https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/DogResnet50Data.npz) bottleneck features
- Inception (https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/DogInceptionV3Data.npz) bottleneck features
- Xception (https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/DogXceptionData.npz) bottleneck features

The files are encoded as such:

```
Dog{network}Data.npz
```

where `{network}`, in the above filename, can be one of `VGG19`, `Resnet50`, `InceptionV3`, or `Xception`. Pick one of the above architectures, download the corresponding bottleneck features, and store the downloaded file in the `bottleneck_features/` folder in the repository.

(IMPLEMENTATION) Obtain Bottleneck Features

In the code block below, extract the bottleneck features corresponding to the train, test, and validation sets by running the following:

```
bottleneck_features = np.load('bottleneck_features/Dog{network}Data.npz')
train_{network} = bottleneck_features['train']
valid_{network} = bottleneck_features['valid']
test_{network} = bottleneck_features['test']
```

In [25]:

```
### TODO: Obtain bottleneck features from another pre-trained CNN.

# ResNet-50 model
bottleneck_features = np.load('bottleneck_features/DogResnet50Data.np
z')
train_Resnet = bottleneck_features['train']
valid_Resnet = bottleneck_features['valid']
test_Resnet = bottleneck_features['test']
```

(IMPLEMENTATION) Model Architecture

Create a CNN to classify dog breed. At the end of your code cell block, summarize the layers of your model by executing the line:

```
Resnet.summary()
```

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: The selected pretrained model is the Resnet-50. This model was developed for image recognition and achieved performance that surpassed the human one. It is therefore suitable for the particular application and it is expected to perform high accuracy. As the basis for the CNN of this particular application, using transfer learning, it is actually very fast to train.

The approach is straightforward. The pretrained Resenet-50 is applied and the last dense layer to match the number of the outputs is attached.

The testing accuracy, shown below is 81%.

In [26]:
```
### TODO: Define your architecture.

# architecture based on Resnet using transfer learning
Resnet_model = Sequential()
Resnet_model.add(GlobalAveragePooling2D(input_shape=train_Resnet.shap
e[1:]))
Resnet_model.add(Dense(133, activation='softmax'))
Resnet_model.summary()
```

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
global_average_pooling2d_3 ( (None, 2048)              0         
_________________________________________________________________
dense_3 (Dense)              (None, 133)               272517    
=================================================================
Total params: 272,517
Trainable params: 272,517
Non-trainable params: 0
_________________________________________________________________
```

(IMPLEMENTATION) Compile the Model

In [27]:
```
### TODO: Compile the model.

Resnet_model.compile(loss='categorical_crossentropy', optimizer='rmsp
rop', metrics=['accuracy'])
```

(IMPLEMENTATION) Train the Model

Train your model in the code cell below. Use model checkpointing to save the model that attains the best validation loss.

You are welcome to augment the training data (https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html), but this is not a requirement.

In [44]:

```
### TODO: Train the model.

checkpointer = ModelCheckpoint(filepath='saved_models/weights.best.Re
snet.hdf5',
                               verbose=1, save_best_only=True)

Resnet_model.fit(train_Resnet, train_targets,
          validation_data=(valid_Resnet, valid_targets),
          epochs=10, batch_size=64, callbacks=[checkpointer],
verbose=1)
```

```
Train on 6680 samples, validate on 835 samples
Epoch 1/10
5952/6680 [=========================>....] - ETA: 0s - loss: 0.1041 -
acc: 0.9716Epoch 00000: val_loss improved from inf to 0.63973, saving
model to saved_models/weights.best.Resnet.hdf5
6680/6680 [==============================] - 0s - loss: 0.1040 - acc:
0.9708 - val_loss: 0.6397 - val_acc: 0.8144
Epoch 2/10
6464/6680 [============================>.] - ETA: 0s - loss: 0.0640 -
acc: 0.9865Epoch 00001: val_loss improved from 0.63973 to 0.59165, sa
ving model to saved_models/weights.best.Resnet.hdf5
6680/6680 [==============================] - 0s - loss: 0.0642 - acc:
0.9864 - val_loss: 0.5916 - val_acc: 0.8084
Epoch 3/10
6464/6680 [============================>.] - ETA: 0s - loss: 0.0415 -
acc: 0.9927Epoch 00002: val_loss did not improve
6680/6680 [==============================] - 0s - loss: 0.0418 - acc:
0.9928 - val_loss: 0.6264 - val_acc: 0.8132
Epoch 4/10
6336/6680 [===========================>..] - ETA: 0s - loss: 0.0317 -
acc: 0.9943Epoch 00003: val_loss did not improve
6680/6680 [==============================] - 0s - loss: 0.0320 - acc:
0.9945 - val_loss: 0.6064 - val_acc: 0.8287
Epoch 5/10
6528/6680 [============================>.] - ETA: 0s - loss: 0.0224 -
acc: 0.9963Epoch 00004: val_loss did not improve
6680/6680 [==============================] - 0s - loss: 0.0224 - acc:
0.9964 - val_loss: 0.6425 - val_acc: 0.8228
Epoch 6/10
6144/6680 [==========================>...] - ETA: 0s - loss: 0.0168 -
acc: 0.9979Epoch 00005: val_loss did not improve
6680/6680 [==============================] - 0s - loss: 0.0175 - acc:
0.9973 - val_loss: 0.6590 - val_acc: 0.8275
Epoch 7/10
5888/6680 [=========================>....] - ETA: 0s - loss: 0.0125 -
acc: 0.9980Epoch 00006: val_loss did not improve
6680/6680 [==============================] - 0s - loss: 0.0127 - acc:
0.9979 - val_loss: 0.6767 - val_acc: 0.8228
Epoch 8/10
5632/6680 [========================>.....] - ETA: 0s - loss: 0.0096 -
acc: 0.9986Epoch 00007: val_loss did not improve
6680/6680 [==============================] - 0s - loss: 0.0104 - acc:
0.9984 - val_loss: 0.6423 - val_acc: 0.8263
Epoch 9/10
6528/6680 [============================>.] - ETA: 0s - loss: 0.0081 -
acc: 0.9989Epoch 00008: val_loss did not improve
6680/6680 [==============================] - 0s - loss: 0.0084 - acc:
0.9990 - val_loss: 0.7046 - val_acc: 0.8120
Epoch 10/10
6656/6680 [============================>.] - ETA: 0s - loss: 0.0077 -
acc: 0.9985Epoch 00009: val_loss did not improve
6680/6680 [==============================] - 0s - loss: 0.0077 - acc:
0.9985 - val_loss: 0.6981 - val_acc: 0.8347
```

```
Out[44]: <keras.callbacks.History at 0x7f1b2afd99b0>
```

(IMPLEMENTATION) Load the Model with the Best Validation Loss

In [45]:
```
### TODO: Load the model weights with the best validation loss.
Resnet_model.load_weights('saved_models/weights.best.Resnet.hdf5')
```

(IMPLEMENTATION) Test the Model

Try out your model on the test dataset of dog images. Ensure that your test accuracy is greater than 60%.

In [46]:
```
### TODO: Calculate classification accuracy on the test dataset.

# get index of predicted dog breed for each image in test set
Resnet_predictions = [np.argmax(Resnet_model.predict(np.expand_dims(f
eature, axis=0))) for feature in test_Resnet]

# report test accuracy
test_accuracy = 100*np.sum(np.array(Resnet_predictions)==np.argmax(te
st_targets, axis=1))/len(Resnet_predictions)
print('Test accuracy: %.4f%%' % test_accuracy)
```

```
Test accuracy: 81.1005%
```

(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.

Similar to the analogous function in Step 5, your function should have three steps:

1. Extract the bottleneck features corresponding to the chosen CNN model.
2. Supply the bottleneck features as input to the model to return the predicted vector. Note that the argmax of this prediction vector gives the index of the predicted dog breed.
3. Use the `dog_names` array defined in Step 0 of this notebook to return the corresponding breed.

The functions to extract the bottleneck features can be found in `extract_bottleneck_features.py`, and they have been imported in an earlier code cell. To obtain the bottleneck features corresponding to your chosen CNN architecture, you need to use the function

```
extract_{network}
```

where `{network}`, in the above filename, should be one of `VGG19`, `Resnet50`, `InceptionV3`, or `Xception`.

In [47]:
```
Resnet_model.summary()
```

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
global_average_pooling2d_3 ( (None, 2048)              0         
_________________________________________________________________
dense_3 (Dense)              (None, 133)               272517    
=================================================================
Total params: 272,517
Trainable params: 272,517
Non-trainable params: 0
_________________________________________________________________
```

In [48]:
```
### TODO: Write a function that takes a path to an image as input
### and returns the dog breed that is predicted by the model.

def Resnet_predict_breed(img_path):
    # extract bottleneck features
    bottleneck_feature = extract_Resnet50(path_to_tensor(img_path))
    # obtain predicted vector
    predicted_vector = Resnet_model.predict(bottleneck_feature)
    # return dog breed that is predicted by the model
    return dog_names[np.argmax(predicted_vector)]
```

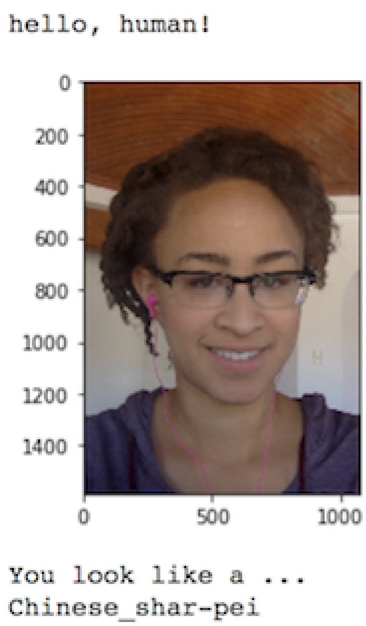
Step 6: Write your Algorithm

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 5 to predict dog breed.

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



(IMPLEMENTATION) Write your Algorithm

In [49]:
```
### TODO: Write your algorithm.
### Feel free to use as many code cells as needed.


def detector(img_path):
    img = cv2.imread(img_path)
    # 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()

    # check for dog or human
    if dog_detector(img_path):
        print("Dog detected of breed",Resnet_predict_breed(img_path))
    elif face_detector(img_path):
        print("Human detected looking like",Resnet_predict_breed(img_
path))
    else:
        print("Unable to detect human or dog ...")
```

Step 7: 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?

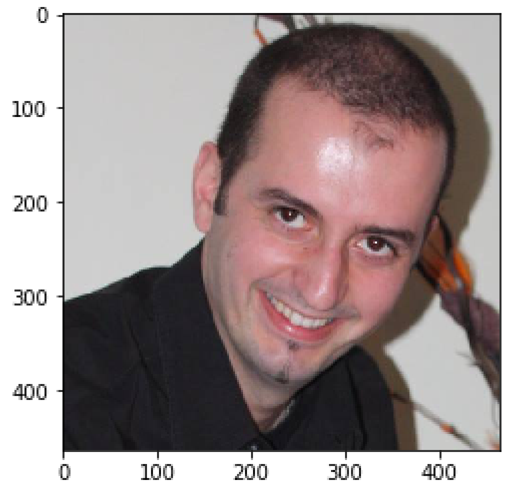
(IMPLEMENTATION) Test Your Algorithm on Sample Images!

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.

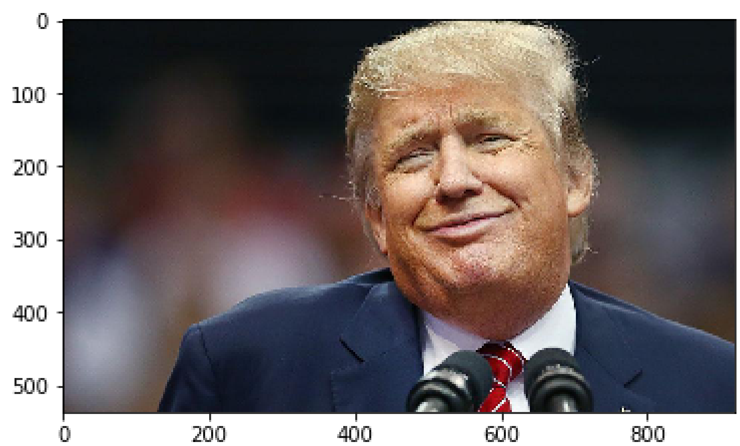
Answer:

In [50]:
```
## 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.
img_name = "images/me.jpg"
detector(img_name)
```



Human detected looking like Xoloitzcuintli

In [51]:
```
img_name = "images/trump.jpg"
detector(img_name)
```

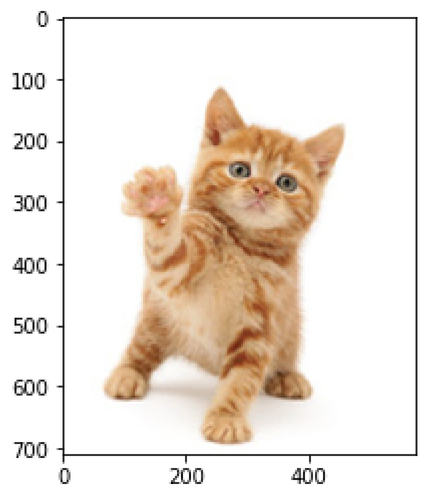


Human detected looking like Lowchen

In [52]:
```
img_name = "images/ted.jpg"
detector(img_name)
```

Unable to detect human or dog ...

In [53]:
```
img_name = "images/cat1.jpg"
detector(img_name)
```

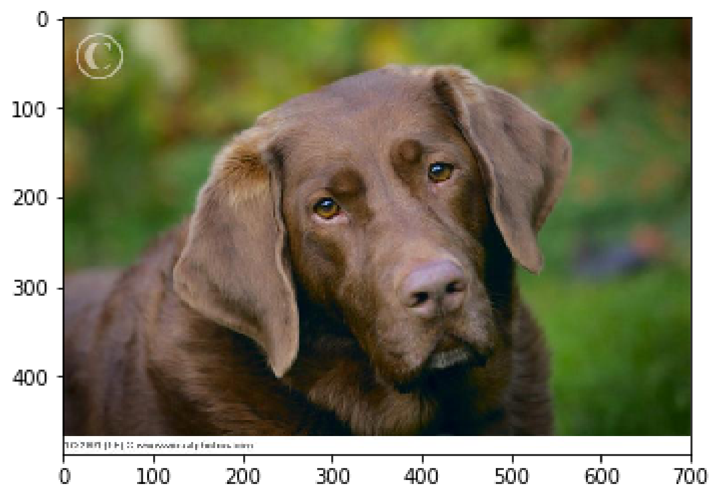


Unable to detect human or dog ...

In [54]:
```
img_name = "images/rottweiler.jpg"
detector(img_name)
```

Dog detected of breed Beauceron

In [55]:
```
img_name = "images/Labrador_retriever_06455.jpg"
detector(img_name)
```



Dog detected of breed Labrador_retriever

Conclusion

The algorithm is performing quite well. In the examples shown above, it is able to succesfully detect dogs and faces as well as ignoring them when they don't exist. Regarding the dog breed, Labrador is succesfully identified whereas the Rottweiler is falsely considered a Beauceron. However, these two breeds bear strong resemblance to each other.

The use of Resnet is giving an accuracy of 81% which given the large number of the outputs is high enough. Moreover, the network itself is relatively small and combined with transfer learning the training is fast.

To improve the algorithm, adding extra layers and increasing the complexity of the model can help to detect more details. More trials using higher number of epochs and image augmentation can also potentially improve the performance.

In [ ]:
```

```