Machine Learning Nanodegree Capstone

Project Proposal: Dog Break Classifier

Deepak Gupta

19th April, 2020

1 Definition

1.1 Project Overview

This project hopes to identify dog breeds from images.This is a fine-grained classification problem: all breeds of Canis lupus familiar is share similar body features and over-all structure, so differentiating between breeds is a difficult problem. Furthermore, there is low inter-breed and highintra-breed variation; in other words, there are relatively few differences between breeds and relatively large differences within breeds, differing in size, shape, and color. In fact, dogs are both the most morphologically and genetically diverse species on Earth. The difficulties of identifying breedsbecause of diversity are compounded by the stylistic dif-ferences of photographs used in the dataset, which featuresdogs of the same breed in a variety of lightings and positions.

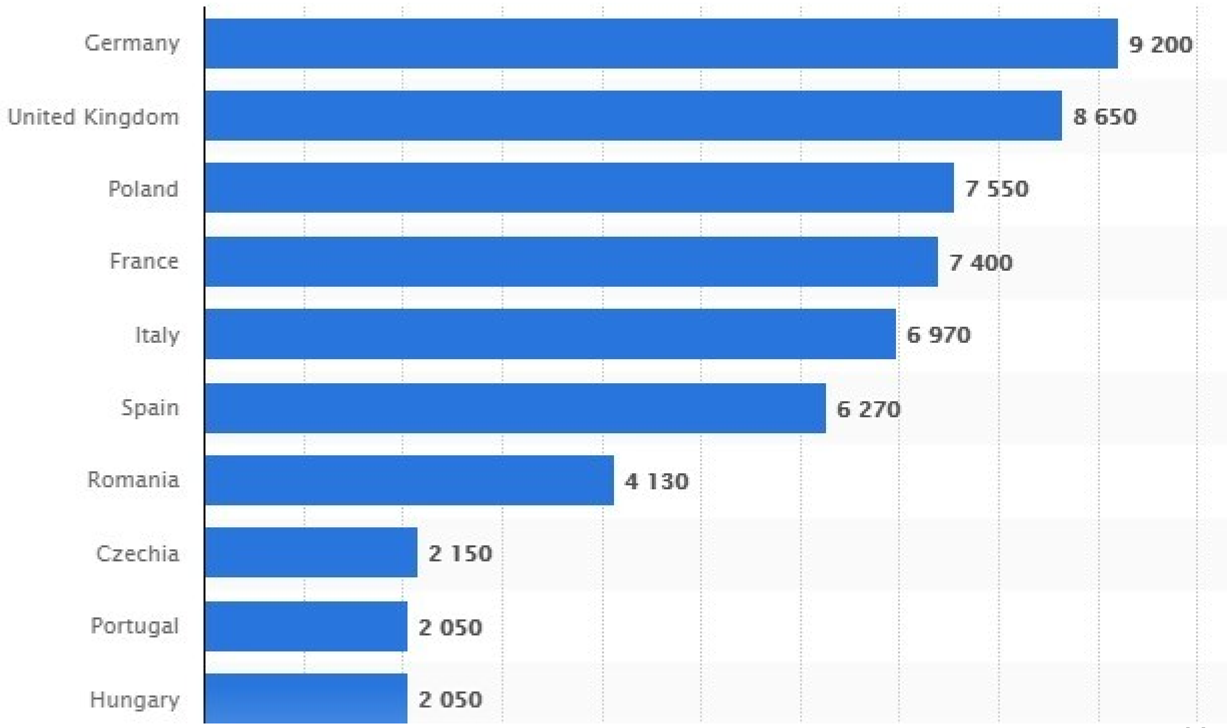


Figure 1: Top 10 most dog-populated European countries in thousands (1000’s)[1]

1.2 Problem Statement

The aim of the project is to build a pipeline to process real-world, user-supplied images. The algorithm will identify an estimate of the dog's breed given an image. When the image is of a human, the algorithm will choose an estimate of a dog breed that resembles the human. If neither a dog or a human is detected, then an error message is output. Therefore, the models in place should be capable of detecting a dog or human in an image, classify the dog to its breed and classify a dog breed that the human resembles.

1.3 Metrics

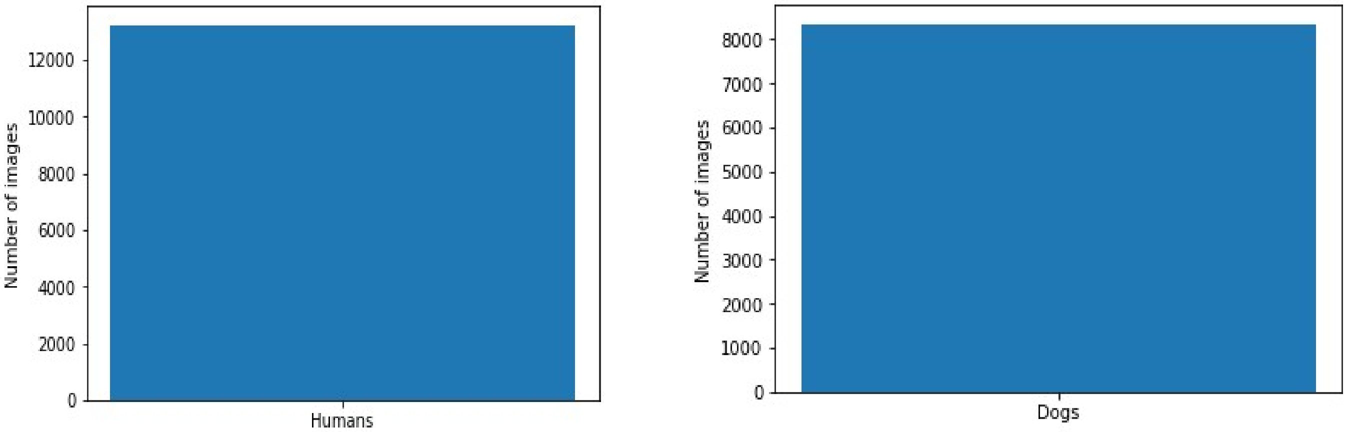
Accuracy will be the main metric used to test both the benchmark model and the solution model. This is so because the problem at hand is a classification task, where the model should classify the images accurately.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

True Positive(TP), True Negative(TN), False Positive(FP) & False Negative(FN).

2 Analysis

The datasets are provided by Udacity. The human dataset[2] contains 13233 images of humans, first names and last names. The dog dataset[3] contains 8351 images of dogs, 133 breeds and each has a representation of 8 images. The images in both datasets have a size of approximately 400 by 400 but will be resized when classifying dog breeds from scratch.



(a) 13233 images in human dataset (b) 8351 images in dog dataset

Figure 2: plots of the number of images in each dataset

The algorithms and techniques are discussed in detail in the methodology section. The benchmark result for the final solution to the problem stattement should surpass 60% test accuracy, a threshold set by Udacity.

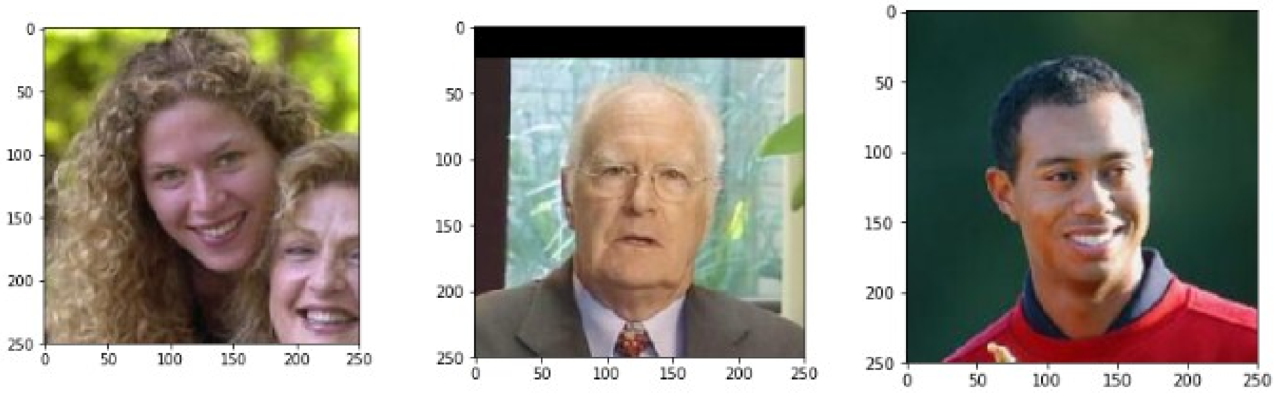


Figure 3: samples of images in human dataset

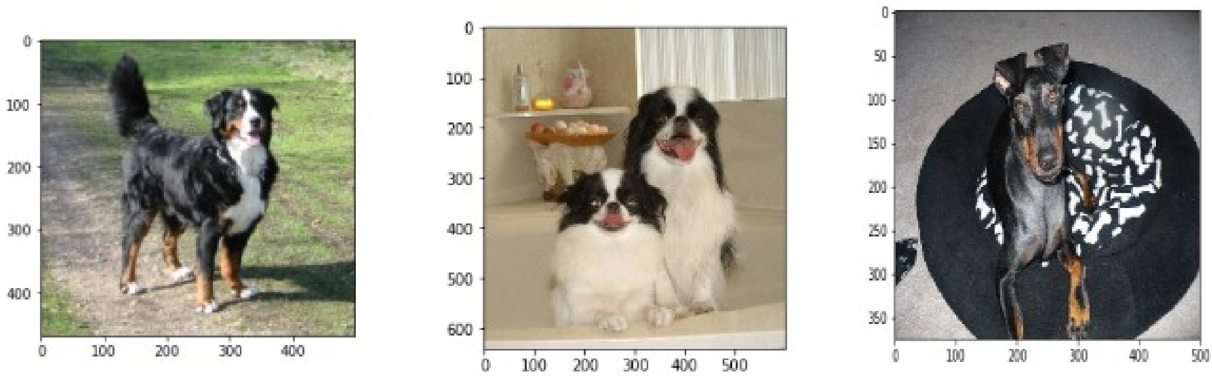


Figure 4: samples of images in dog dataset

3 Methodology

3.1 Detecting Humans

Open CV's implementation of Haar feature-based-cascade-classifiers is used to detect human faces in the user-supplied images. There are many pre-trained detectors given by Open CV, for this project they are stored in the 'haarcascades' directory. The images are converted to greyscale before being passed to a face detector. The face detector is tested with 100 images each from the human and dog datasets. The face detector finds 98% of the human images have a human face and 23% of the dog images have a human face.

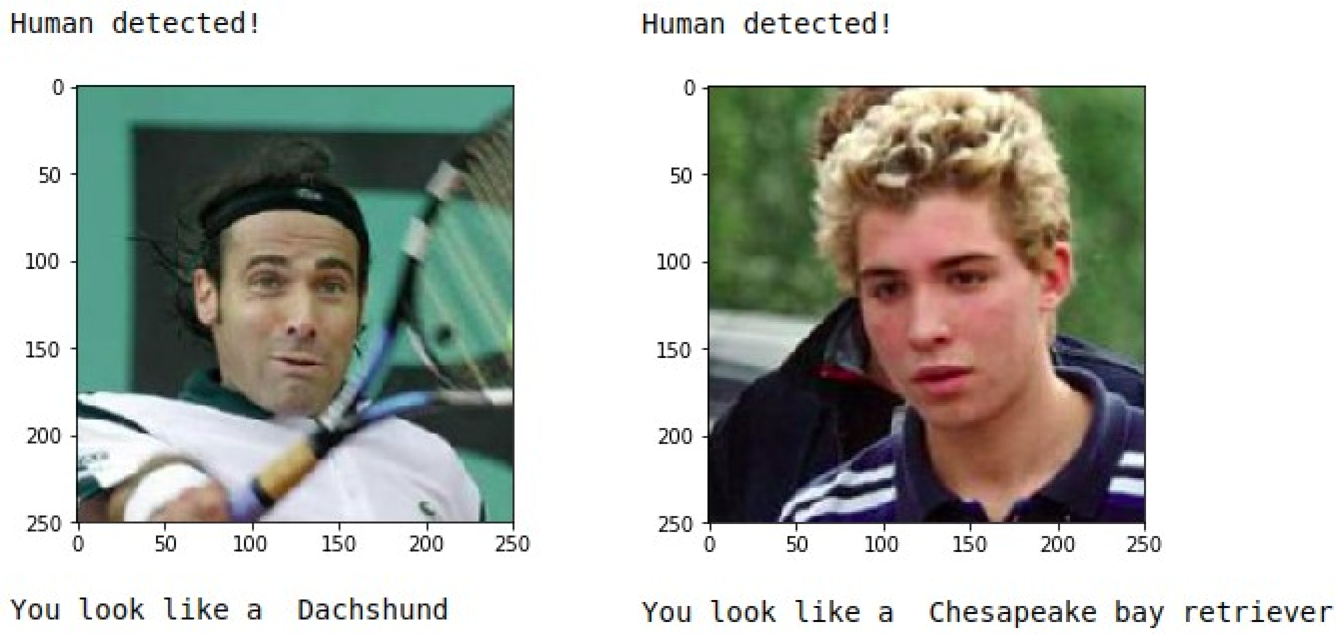


Figure 5: Human face detector results on two images from the human images database

3.1 Detecting Dogs

A pre-trained VGG-16 model is used to detect dogs in images. It is used with weights trained on ImageNet, a large and popular dataset used for image classification tasks. ImageNet has more than 10 million URLs where each URL links to an image containing an object from one of 1000 categories. Dog breeds occur consecutively on a dictionary from ImageNet from keys 151 to 268 inclusively, that is, from ’Chihuahua’ to ’Mexican hairless’. Therefore, the VGG-16 model is excepted to return an index betwwen 151 to 268 (inclusive).

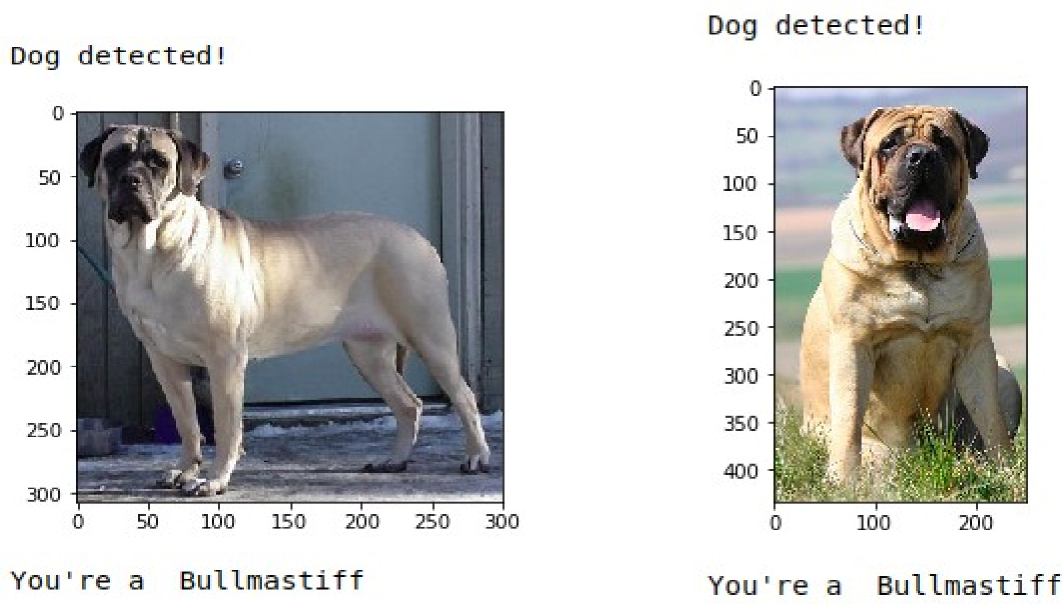


Figure 6: Dog detector results on two images from the human images database

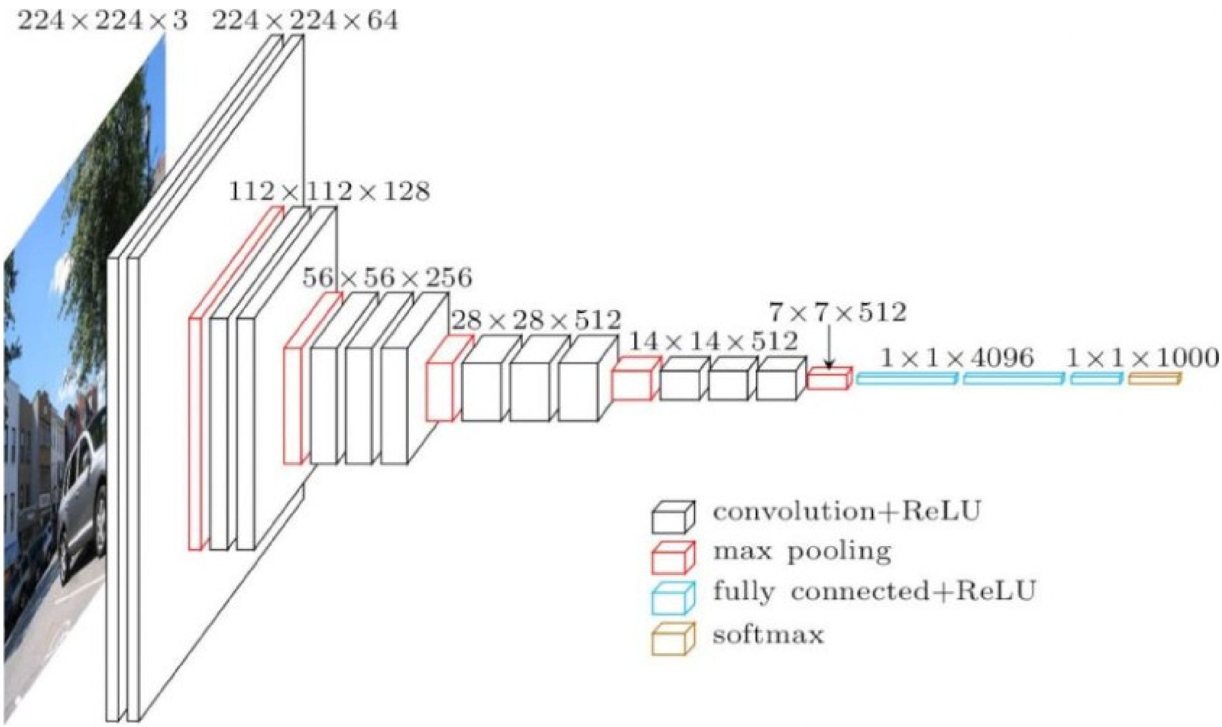


Figure 7: VGG-16 model architecture [4]

3.3 Classifying Dog Breeds From Scratch

We have determined how to detect humans and dogs in images in the last sections using Haare feature-based-cascade classifiers and a VGG-16 model respectively. Now, a convolutional neural network (CNN) will be built from scratch with the aim of classyfing dog breeds accordingly with the problem statement. However, the images in the datasets have to be preprocessed before passing them as into the CNN model. The images are resized to 224 by 224 because this is the default for ResNet [5] (which will be used for transfer learning later on). Data augumentation is done through flips, rotations and crops. The dataset is split into train data (for training the model), validation data (for choosing the most accurate model), test data (for testing the model). CNN layers are utilized for feature extraction and image generalization. The dropout layer ensures that overfitting is dealt with and the linear classifier converts the extracted features to a classified type.

3.4 Classifying Dog Breeds With Transfer Learning

The CNN model built from scratch can be improved significantly from its 11% test accuracy. This can be done through transfer learning. The CNN model will now use ResNet50 with the preprocessed data for the previous CNN model.

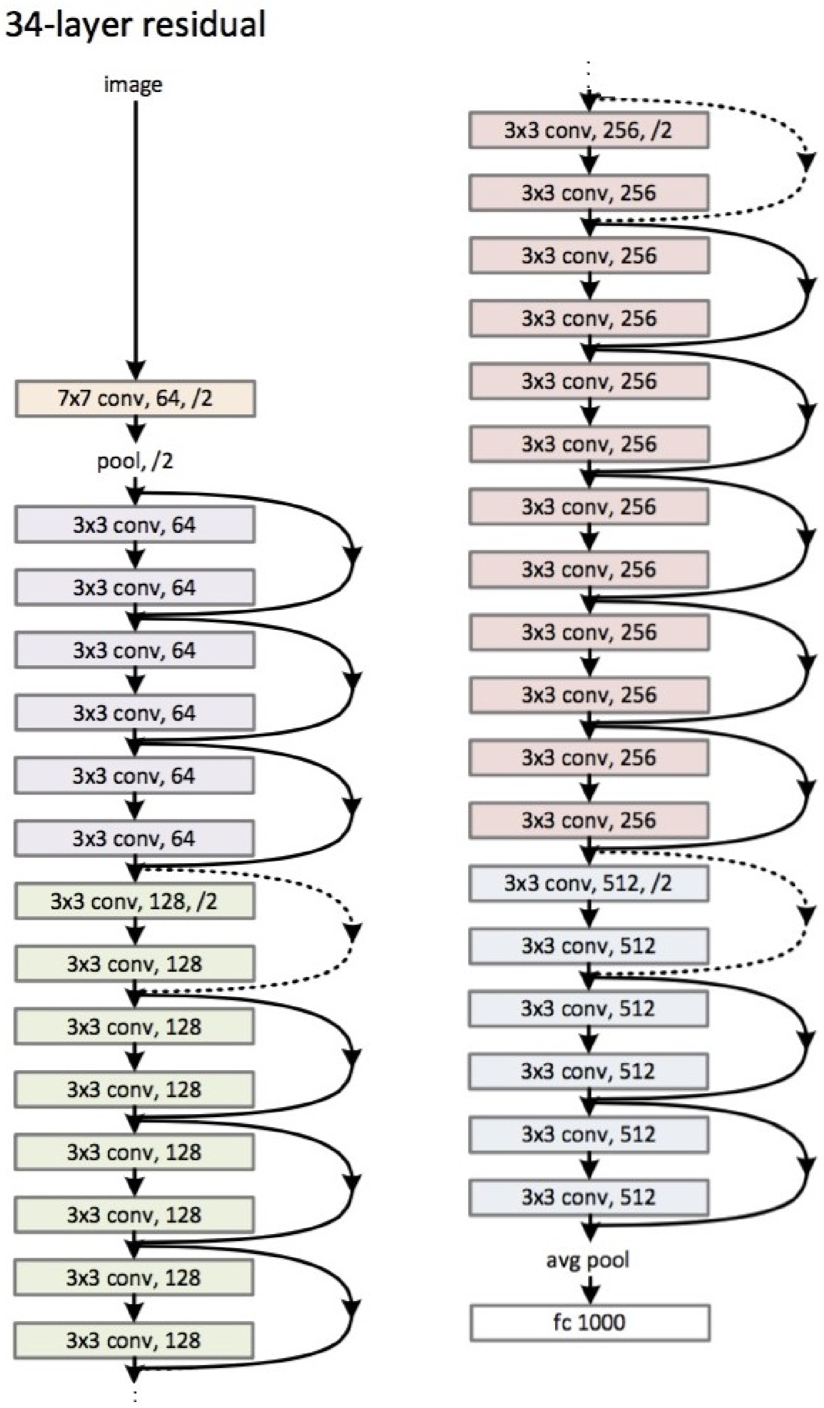


Figure 8: Multi-layer ResNet50 architecture [6]

4 Results

The final CNN model created with ResNet50 had a test loss of 2.630944 and a test accuracy of 65% (546/836). We look at 10 samples from both datasets and their results from the model.

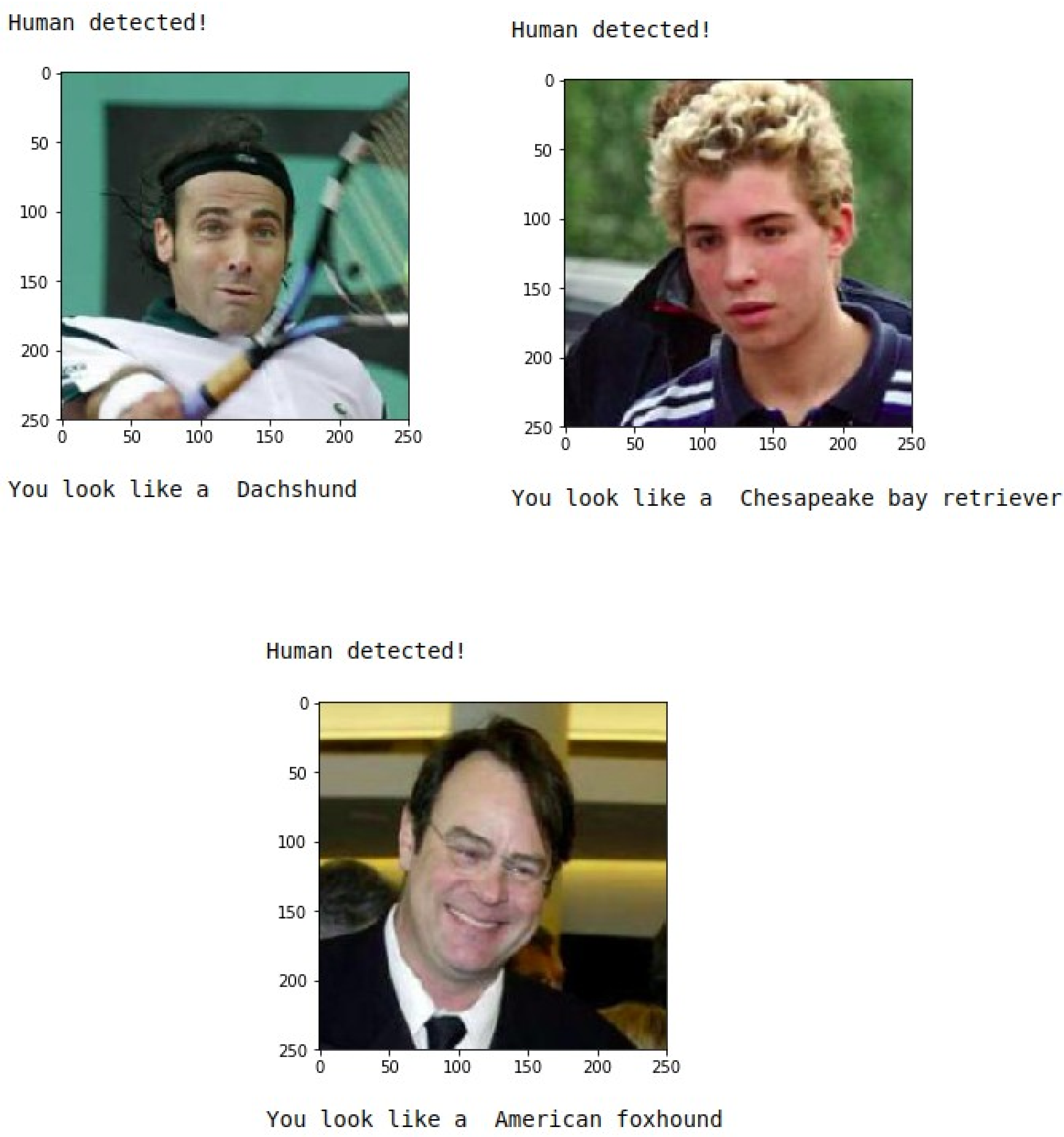
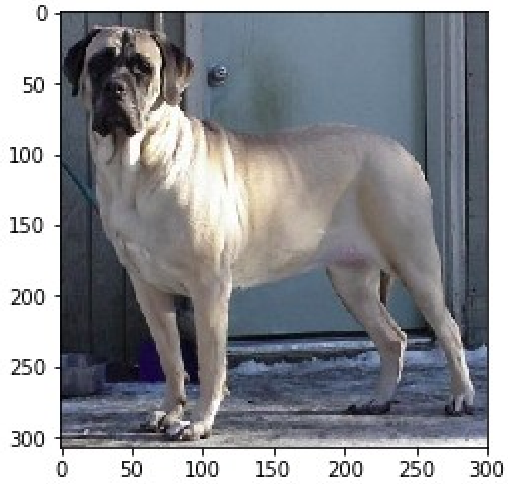


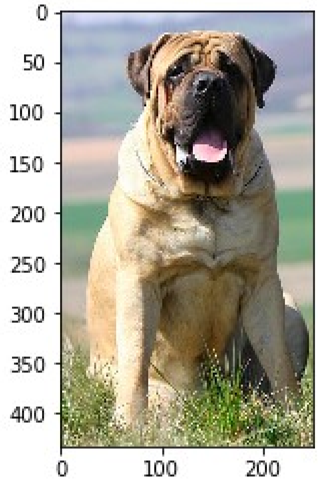
Figure 9: Results from human dataset

Dog detected!



You're a Bullmastiff

Dog detected!



You're a Bullmastiff

Dog detected!

You're a Bullmastiff

Figure 10: Results from dog dataset

5 Future Work

Future work should further explore the potential of convolutional neural networks (CNN) in dog breed prediction. Given the success of our keypoint detection network, this is a promising technique for future projects. That said, neural networks take an enormous time to train and wewere unable to perform many iterations on our techniquedue to time constraints. We recommend further exploration into neural networks for keypoint detection, specifically by training networks with a different architecture and batch iterator to see what approaches might have greater success. Also, given our success with neural networks and keypoint detection, we recommend implementing aneural network for breed classification as well since thishas not been performed in the literature. We were unable toexperiment with this approach due to the time constraintsof neural networks but believe that they would match ifnot improve upon our classification results. Ultimately,neural networks are time consuming to train and iterateupon, which should be kept in consideration for futureefforts; still, neural networks are formidable classifiers that will increase prediction accuracy over more traditional techniques.[7]

6 Bibliography

[1] Statistica. Available at: https://www.statista.com/statistics/414956/dog-population-european-union-eu-by-country/ .

[2] Udacity, Human Dataset. Available at: https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/lfw.zip/ .

[3] Udacity, Dog Dataset. Available at: https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/dogImages.zip/ .

[4] Neurohive. Available at: https://neurohive.io/en/popular-networks/vgg16/ .

[5] Keras Documentation. Available at: https://keras.io/applications/#resnet/ .

[6] Stack Exchange Available at: https://i.stack.imgur.com/XTo6Q.png/

[7] Stanford. Available at: https://web.stanford.edu/class/cs231a/prev_projects_2016/output%20(1).pdf.