Building a Dog Breed Classifier

Udacity Machine Learning Engineer Nanodegree

Capstone Proposal

ChihYu Yeh

Domain Background

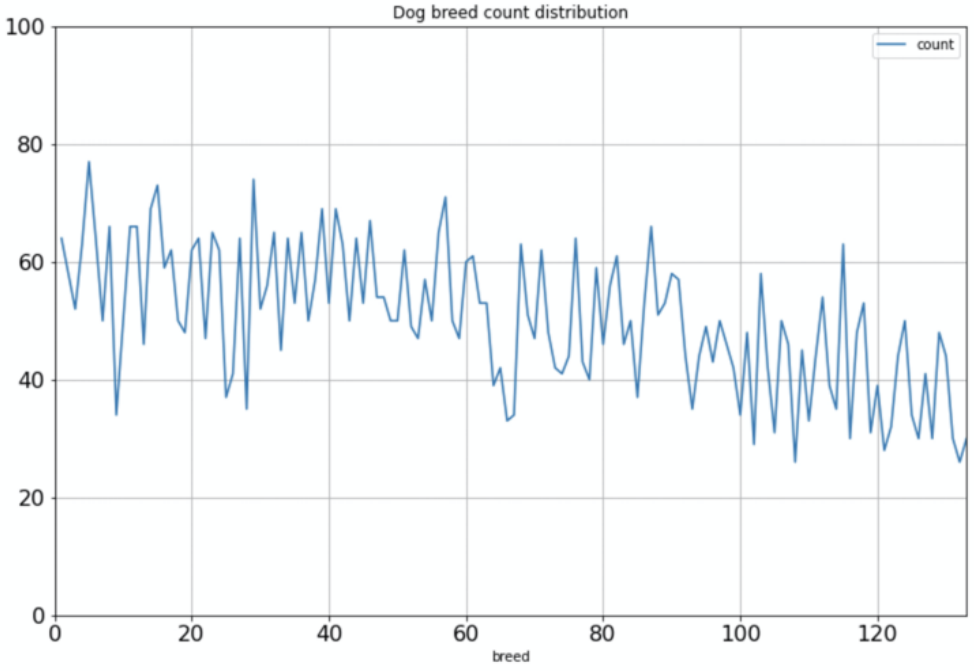
Building a dog breed classifier is a computer vision task. Since the ImageNet challenge, image classification task performance using deep learning techniques becomes better and better and even outperforms human performance. Besides better performant neural network architectures coming out each year, one special technique called transfer learning really makes applying deep learning to real life much easier. With transfer learning, we can take one predefined architecture with pretrained weight, and fine-tune only last few layers of the model to quickly bring the new task performance to a quite accurate level.

Problem Statement

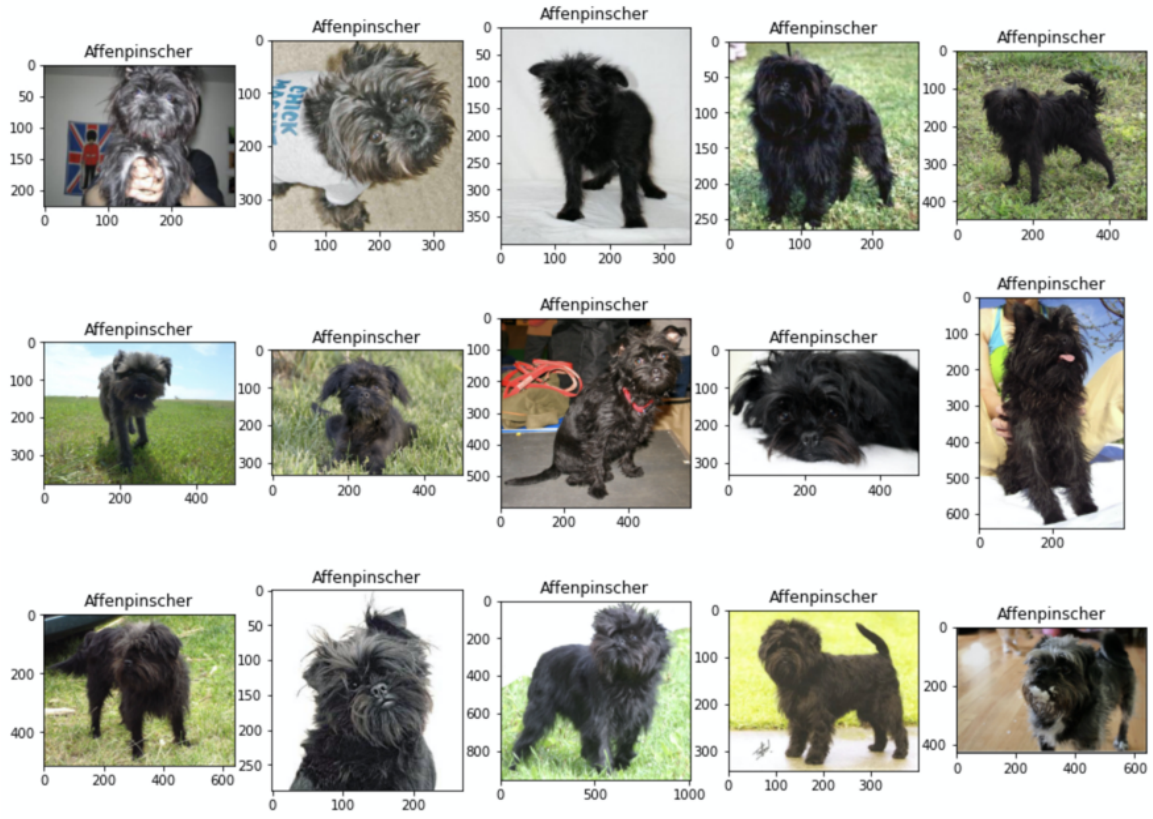
In this project, I will apply the transfer learning technique to solve the dog breed classification task. This is a multi-classification problem, since there are multiple dog breeds out there. Besides classifying dog breeds, if a human is inside an input image, we will predict a dog breed that is the most similar to that human. Otherwise, the output will print out a message that indicates error.

Datasets and Inputs

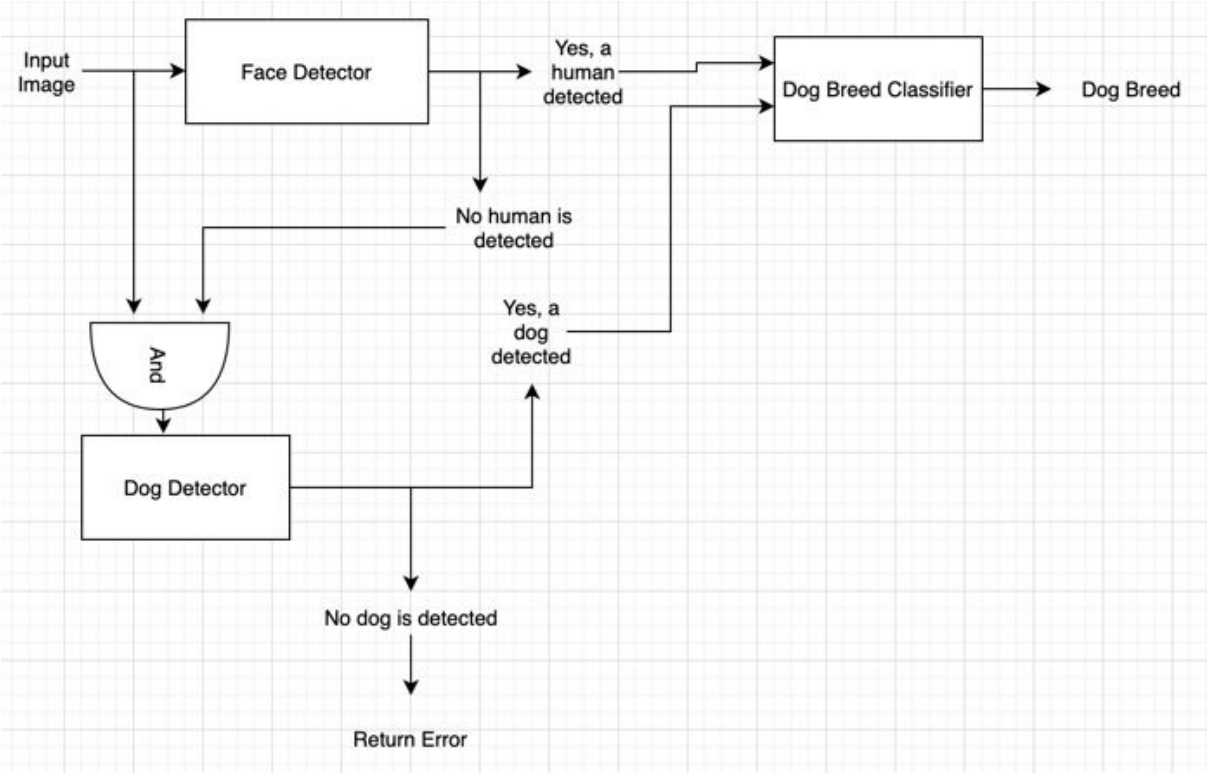
Datasets contain both dog and human images. There are 13223 human images and 8351 dog images(133 dog breeds). Also in dog images, we have training/validation/testing datasets. The image below is the dog breed distribution for the training data



Sample input images and respective image sizes are shown below:



Solution Statement



- Face Detector: OpenCV's implementation of Haar feature-based cascade classifiers to detect human faces in images. Pretrained models can be accessed on github.
- Dog Detector: A pretrained VGG-16 model, and 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.
- Dog Breed Classifier: A large-scale pretrained model for computer vision, called Big Transfer(BiT). Google releases the BiT-S and BiT-M pretrained models, which are trained with ILSVRC-2012 (1.28M images with 1000 classes), ImageNet-21k (14M images with ~21k classes) respectively.

Benchmark Model

The key ingredient of the project is the dog breed classifier, so I will use the same components for the face and dog detector in the benchmark model and my proposed model. The face detector is based on an OpenCV algorithm. Finally, for the dog breed classifier, I will design a simple(naive) solution myself to make the benchmark model. It should be a simple CNN neural network model.

Evaluation Metrics

The evaluation metrics is the dog breed prediction accuracy given test images. However, if the data is severely unbalanced, then we should look at other metrics, such as recall, precision, f1, etc.

Project Design

The overall workflow is explained as the following:

- Data exploration
    - Check number of training data, validation data, and testing data.
    - Draw target(dog breed) number distribution based on training data to check if they are severely unbalanced. If so, try add more data or use sampling techniques to preprocess the training data.
- Preprocessing
    - Make sure the image size would be the same for all input data to neural network models
    - Apply normalization to image data for neural network models
    - Try to apply the same preprosessing step as the pretrained model does
    - Try data augmentation if possible to enhance model stability(generalization ability)
- Training and testing the three models
    - Face detector(using one pretrained CV algorithm)
    - Dog detector(using pretrained VGG-16 model)
    - Dog breed classifier(using pretrained Big Transfer(BiT) model)
- Refinement
    - Apply transfer learning using a pretrained Big Transfer model, then fine-tune the model hyperparameters(batchsize, optimizer, learning rate, model architecture) to suit the project's need
- Performance evaluation
    - Use dog breed classification accuracy as the performance metric