Udacity Capstone Proposal: Dog Breed Classifier

This is my capstone proposal for the Machine Learning Engineer Nanodegree. This proposal paper follows the following structure to give an overview over the complete project.

First, I will describe the domain background, in other words the field of research where the project is derived. Next, I will formulate a problem statement for which a solution will be defined. Afterwards, I will give a short summary over the datasets and inputs which will be used in this project. In the next step I will propose a solution statement, where I will describe my thoughts on how to solve the defined problem. A benchmark model will be introduced, so the solution can be compared to certain performance measures and A set of evaluation metrics will be introduced, that is, functional representations for how the solution can be measured. Finally, I will present a project design and describe the steps necessary to develop a solution and obtain results.

Domain Background

The task I chose to tackle is a traditional task in machine learning. The classification of images. There have been many different approaches to gain a high accuracy. In this project, I will try different approaches to gain a high accuracy on this task.

Problem Statement

So the underlying problem is to determine whether or not we have a human or dog in our picture. If it is a human, we want our model to recognize that and then find out which dog breed this human resembles most. If our model is confidant it found a dog, we want to have the probabilities about the different breeds this dog could belong to.

Datasets and Inputs

As inputs we got a human and a dog dataset, which consist of 13233 and 8351 pictures, respectively.

Project Design

The project design has the following outline:

Step 0: Import Datasets

Here, I will load the aforementioned dog and human datasets and import the necessary libraries as well.

Step 1: Detect Humans

I first created a human detector. A pretrained face detector will be used for this purpose ('haarcascades/haarcascade_frontalface_alt.xml').

Step 2: Detect Dogs

Next, I will use a pre-trained VGG-16 model for the dog detector. It has to be considered, that the input images must be cropped to 244x244 pixel tensors in order to be used be the model.

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

In this step I will split our datasets into train, validation and test sets, extract the possible class names and define a model architecture which is inspired by the VGG 16 Model, developed by Oxford University. The goal is to attain a test accuracy of at least 10%.

Step 4: Create a CNN to Classify Dog Breeds (using Transfer Learning)

In this step I will use a pre-trained ResNet50-Model for transfer learning, which has proven to be very powerful for a multiclass classification task like ours before. My CNN must attain at least 60% accuracy on the test set.

Step 5: Write your Algorithm

Next, I will write my own algorithm that accepts a file path to an image and first determines whether the image contains a human, dog, or neither. This is the conditions my algorithm will consider during its task:

- 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.

Step 6: Test Your Algorithm

In a last step I will test my algorithm on sample data which I have downloaded separately. By doing so I will get a direct feedback on my work and can conclude on where possible weaknesses can be identified.

Evaluation Metrics and Benchmark Model

As evaluation metrics we have predefined accuracy scores of 10% and 60%, respectively. In addition, I will observe the loss and cross validation loss during the training process of our models in order to identify whether our models overfit the training data or hyperparameters have to be finetuned in order to attain a higher accuracy.

Solution Statement

The solution statement will state if I am content with the model accuracy and the output of sample images I. Based on this, I will define points where our model can be improved.