

Convolutional Neural Network for Dog Breed Classification

Udacity Machine Learning Engineer Nanodegree Capstone Project

Written by **Alex Augusto Costa Machado** on **December 10th 2021**

1 Definition

1.1 Project Overview

For computer vision tasks, such as a dog breed classification, an algorithm mostly used is a Convolutional Neural Network, or CNN for short, which is a model that extracts features, like texture and edges, from spatial data.

The history behind this algorithm started during the 1950s, but it is around the year 2012 that CNNs saw a huge surge in popularity after a CNN called AlexNet achieved state-of-the-art performance labeling pictures in the [ImageNet](#) challenge. Alex Krizhevsky et al. published the paper "ImageNet Classification with Deep Convolutional Neural Networks" describing the winning AlexNet model.

1.2 Problem Statement

The purpose of this project is to use a Convolutional Neural Network to classify dog breeds using images as input. For images that contain a human instead of a dog, the algorithm should display which dog breed resembles the human in the picture.

1.3 Metrics

The metric to evaluate the quality of the classifier is the accuracy of the dog breed predictions.

2 Analysis

2.1 Data Exploration

There are two datasets for this project: one is a dog dataset containing 8351 images of 133 different dog breeds, and the other is a human dataset consisting of 13233 photos of 5749 different people. The dog dataset is divided in train, validation and test datasets.

2.2 Exploratory Visualization

2.3 Algorithms and Techniques

Three main algorithms will be used for this project: a human face detector, a dog detector, and a dog breed classifier. More about each of these algorithms is described below.

- Human Face Detector: an OpenCV's implementation of Haar feature-based cascade classifier to detect human faces in images. The pre-trained face detector is stored as XML file on [GitHub](#).

- Dog Detector: a pre-trained VGG-16 model with weights that have been trained on [ImageNet](#), a very popular dataset used for image classification and other vision tasks.
- Dog Breed Classifier: a pre-trained VGG-11 model on which the last layer is altered to output 133 classes and trained further using the dog dataset mentioned earlier.

2.4 Benchmark

The benchmark for the dog breed classifier will be a simple CNN model designed from scratch that will be compared to the final CNN obtained from transfer learning using a pre-trained VGG-11 model.

3 Methodology

3.1 Data Preprocessing

3.2 Implementation

The workflow for this project is as follows:

- Data exploration: verify the information contained in the datasets.
- Preprocessing: crop images to make sure all have the same size and normalize the input data.
- Training and testing: create algorithms to train all three models mentioned earlier.
- Refinement: use transfer learning to improve accuracy of dog breed classifier.
- Performance evaluation: observe classification accuracy as the performance metric.

3.3 Refinement

4 Results

4.1 Model Evaluation and Validation

4.2 Justification

5 Conclusion

5.1 Free-Form Visualization

5.2 Reflection

5.3 Improvement