Convolutional Neural Network Project

Dog Breed Classifier

Han Chao Zhao

November 12, 2020

Domain Background

Convolutional neural network (CNN) is a deep learning algorithm that is well known for its effectiveness in analyzing images. Being able to identify a dog breed has been a popular problem in the machine learning space along with other type of classifiers like plant identification. For this project we will try to once again use CNN to solve the dog breed classifier problem.

Problem Statement

There are two problems we wish to solve here. First we want the user to be able to identify a dog's breed with its picture. Second, we want the user to be able to identify what dog breed a human looks like by uploading a picture of the said human.

Datasets and Inputs

The project will use 8351 dog images separated into training (6680), validation (835), and test (836) datasets. All the dog pictures will be resized to 224x224 and normalized. There are a total of 133 type of breed included in the dog data.

There are total 13233 human images that will be used for human identification and to classify which dog breed a human resembles the most.

Solution Statement

For determine whether a picture is human, we use OpenCV's implementation of Haar feature-based cascade classifiers. For dog detection we used a pre-trained VGG-16 model for the job. First a CNN model will be built from scratch to achieve more than 10% accuracy when identifying dog breed. Afterwards, we will use transfer learning to create a CNN model that should have more than 60% accuracy as our final tool for the classifier.

Benchmark Model

Based on the input we have, a random guess would provide an accuracy of 1/133 which is less than 1%. The CNN model we build from scratch should achieve a better accuracy than random guessing.

Evaluation Metrics

Accuracy will be used as to evaluate the models. The CNN model built from scratch should obtain at least 10% accuracy. The CNN model using transfer learning should achieve at least 60% accuracy.

Project Design

Step 1:

Data import for human and dog images.

Step 2:

Using OpenCV's implementation of Harr feature-based cascade classifiers to determine whether a given picture is a human.

Step3:

Using a pre-trained VGG-16 model to detect whether a given picture is a dog.

Step4:

Creating a CNN model from scratch to classify dog breeds. The test accuracy should be greater than 10%.

Step5:

Creating a CNN model using transfer learning and the test accuracy should now increase to be at least 60%.

Step 6:

Use the model created using transfer learning to classify dog breed for dog pictures and determine which dog breed a human looks like for human pictures.