Capstone Proposal

Jorge Ramírez Carrasco

Dog Breed Classifier (CNN)

Domain Background

Image classification is one of the most important visual tasks in many fields. The idea is to categorize images between a predefined number of classes.

To solve the problem, it is required to apply computer vision techniques and machine learning where convolutional neural networks have shown a great performance. The relationship between these tasks and the convolutional neural networks started from the late 1980s[1], but due to the evolution in computing power they have become popular in the last decade when we have seen an increase in the usage in the industry and academia. Convolutional neural networks are feedforward networks composed mainly by convolutional layers, pooling layers and fully connected layers[2].

The personal motivation about this project comes from the interest of solving image classification using a custom dataset and the knowledge learnt from the course to solve these problems in Amazon Web Services with GPUs.

Problem Statement

The problem of the proposal is to categorize dog and human images estimating the closer canine's breed. Given an image of a dog, the algorithm will identify an estimate of the canine’s breed. If supplied an image of a human, the code will identify the resembling dog breed. From the initial dataset a subset of images will be stored for testing in order to measure the estimations' quality.

Datasets and Inputs

There are two datasets included in the project:

- **Dog Dataset:** It contains 8351 dog images structured in 133 folders by dog's breed class. The dataset is well distributed between the 133 classes. The sizes are different with a standard resolution. Downloaded from Udacity AWS S3. The purpose of the dataset is to train and test a convolutional neural network that classifies dog's breed.

- **Human Dataset:** It contains 13233 human images structured in folders by celebrities' name. The sizes are different with a standard resolution. Downloaded from Udacity AWS S3. The purpose of the dataset is to detect human faces and test the dog's breed convolutional neural network.

---

[1] (Leonidas Lampropoulos, 2016)
[2] (Wang, 2017)

Solution Statement

The solution is composed of several steps. The main idea is to create a classifier able to estimate the dog's breed from an image and deploy it in a web application.

First the datasets will be analyzed, then the purpose will be to create a dog detector using a pre-trained VGG-16 model and a human detector using Haar feature-based cascade classifiers.

The detectors are a key step because the system will provide an estimate of the dog's breed if a dog is detected, an estimate of the dog's breed that is most resembling if a human is detected and it will flag an error otherwise.

To perform the estimation of dog's breed it is required to train and test a convolutional neural network able to input a detected dog from an image and outputs the canine's breed. Initially the convolutional neural network will be trained from scratch and then using transfer learning starting with a pre-trained VGG-16 model from the dataset ImageNet

Finally, the model and the detectors will be deployed and connected with a web app that will allow users to test and interact with the system.

Benchmark Model

The convolutional neural network will be compared against other classification methods like logistic regression and support vector machines.

Additionally, the performance will be compared between the following convolutional neural networks: pre-trained model with ImageNet, dog breed's classifier trained from scratch and dog breed's classifier trained with transfer learning.

Evaluation Metrics

Model's performance will be compared using a classical metric for multiclass image classifier like accuracy. It is also the proposed metric in the project documentation. Accuracy's formula is:

$$Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made}$$

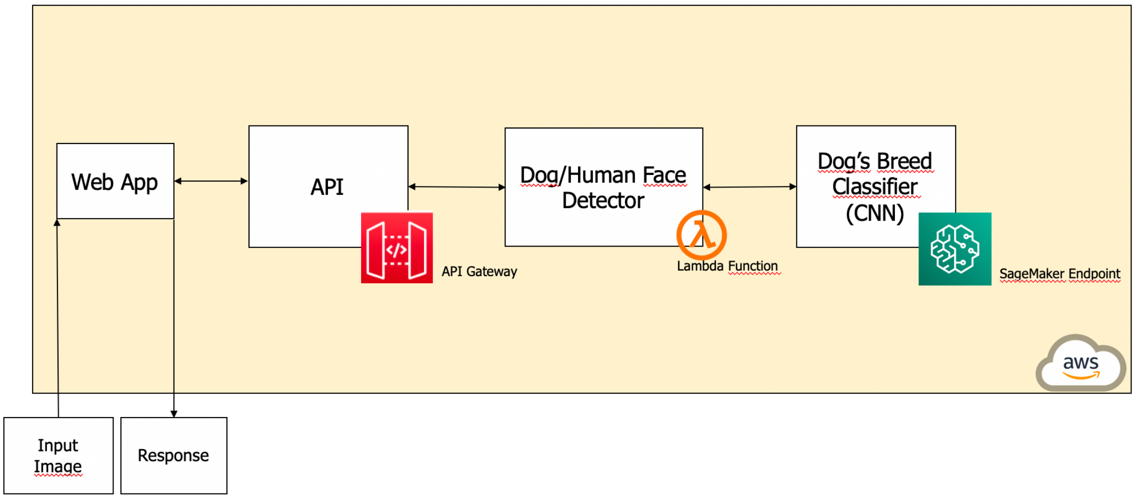
It works well only in cases where the classes are not unbalance where precision or recall will be a better metric.

Project Design

The project is composed by two workflows. First, a model workflow related with data analysis and processing, implementation of algorithms and evaluation that will end in the deployment of the model in a SageMaker Endpoint. Second, a workflow related with the app workflow including the deployment of the required components to make the classifier available through a web app.

The model workflow starts with data analysis and processing of the images to fully understand the dataset features. Then, it is required to work in transformers for the data loader and in the model architecture. These transformers will include a resizing of the images and detectors. The model training is performed once the data loader is ready and the dataset is split in training, validation and testing datasets. Once the model is trained, validated, tested and compare between the different algorithms of the benchmark it can be deployed in a SageMaker Endpoint.

The app workflow is summarized in the following diagram:



As described in the solution statement the system is composed by a web application that is connected through an API gateway with a lambda function that detects human faces and dogs and uses the SageMaker endpoint connected with the trained convolutional neural network that estimates dog's breed.

The flow starts when the user inputs an image and it ends when the system provides the response message. All the components are deployed in Amazon Web Services.

References

Leonidas Lampropoulos, A. C. (2016). Neural network image classifier, Google Patents. https://patents.google.com/patent/US10007866B2/en

Wang, W. R. (2017). Review, Deep Convolutional Neural Networks for Image Classification: A Comprehensive, The MIT PressJournals. https://www.mitpressjournals.org/doi/full/10.1162/neco_a_00990