Machine Learning Nanodegree Capstone Project Proposal

Dog Breed Identification

Daniel Fernando García Rodríguez

Domain Background

Automatic image classification is one of the most important areas of machine learning, both in research and commercial applications. Since 2012 Convolutional neural networks (CNN) based image classification techniques dramatically improved the then current benchmarks setting CNN as SOTA for computer vision while propelling the hype of the so called “Deep Learning revolution”. The goal of this project is to explore the CNN-based approach for image classification using a dog breed dataset.

Problem Statement

Although former computer vision approaches were certainly capable of some image classification task, Deep Learning approaches just outperformed them. Among image classification problems animal breed differentiation is a challenging one because the differences can be subtle and complex; because of that applying deep learning could be an effective solution. The aim of this project is to develop a model capable of distinguishing between dog breeds when the user submits a dog image. This process implies two steps: recognizing if the picture is from an actual dog, and if it is then identifying its breed. This supervised learning approach requires a labeled dataset of dog pictures with its corresponding breed.

Dataset and Inputs

The provided dataset consists in a set of dog pictures labeled with its corresponding breed; it also includes a series of human pictures. The dataset is divided in two separated files, one for dogs and one for humans, both hosted by Udacity. The human set is required for the Haar cascade part of the project which consist in differentiating between whether there is a dog in a picture; then a neural network model will differentiate the breed of a dog picture supplied as input.

The dataset is divided into 3 segments: Train consisting in between 30 – 70 pictures per breed, Test segment with 6 – 10 pics per breed and Validation also 6 – 10 per breed; all containing the same 133 dog breeds. The human dataset contains 5794 pictures of (you guess it!) humans. Not all the dog pics contains only dogs.

Solution Statement

The proposed solution consists in a deep learning-based approach using CNN to detect dog breeds from the supplied picture. The accuracy rate is expected to be around 80%. The model will differentiate if the supplied picture presents a pure breed or mixed breed dogs, also should be capable of recognizing humans. Multiclass logarithm loss will be used as performance metric.

Benchmark Model

As the same problem with similar dataset is present in Kaggle, we can use its leaderboard as benchmark. Nevertheless, top performers report loss scores lower than 0.01 which suggest overfitting. I'll stick to an 80% accuracy score considering time and data constrains.

Evaluation Metrics

Multiclass class log loss, which measures the performance of a multiple class classification model where the prediction score is a probability (between 0 and 1). In this project the output consists of a 133 sized probabilities vector assessing how likely the input image belongs to each one of the 133 breeds.

The training method goal is to minimize the multi class log loss. The other metric is the prediction's accuracy, there are other metrics like sensibility or f-score, but I didn't select **References**
1. Paul Viola and Michael J. Jones. Robust real-time face detection. International Journal of Computer Vision, 57(2):137–154, 2004.
2. Rainer Lienhart and Jochen Maydt. An extended set of haar-like features for rapid object detection. In Image Processing. 2002. Proceedings. 2002 International Conference on, volume 1, pages I–900. IEEE, 2002 them because I don't find them relevant for the project's requirements.

Project Design

Following template's steps, the project will consist of:

1. Importing Datasets.
2. Developing Haar Cascade classifier to Detect Humans.
3. Implementing pre-trained VGG-16 model to Detect Dogs.
4. Developing a full detection classifier from scratch
5. Constructing a full detection classifier using a pretrained model with transfer learning\
6. Testing

References

1. Paul Viola and Michael J. Jones. Robust real-time face detection. International Journal of Computer Vision, 57(2):137–154, 2004.

2. Rainer Lienhart and Jochen Maydt. An extended set of haar-like features for rapid object detection. In Image Processing. 2002. Proceedings. 2002 International Conference on, volume 1, pages I–900. IEEE, 2002