DOG BREED IDENTIFICATION

Machine Learning Scientist Capstone Project Proposal

Chaoyi Wang

Domain Background

Automatic image classification is one of the major topics in the research and application field. This category of techniques provides essential component for applications such as auto-pilot systems, product quality control, retail, radiography, other scientific research, etc. Automatic image classification is commonly achieved by convolutional neural network (CNN), a deep learning framework.
The motivation for proceeding with this project is to learn and practice using convolutional neural network in image classification, starting with dog breeds. The ultimate goal is to develop a modified deep learning model for interpreting seismic reflection maps for auto identification of natural or artificial features.

Problem Statement

Dogs are people's best friends. However, there is literally countless number of dog breeds in the world. How can one tell what the breed of a dog is if first met or given a picture/video clip? This project aims to develop a deep leaning model that can distinguish a breed of a dog if a picture of the dog is supplied by the user. The finished model should firstly feature the ability to distinguish whether the supplied picture is a dog or not. Secondly, the model should accurately identify pure-breed dog. Additionally, the model should exhibit good performance on predicting dogs with mixed breeds.

Datasets and Inputs

The dataset is supplied by Udacity machine learning nanodegree program. Additional dataset is found in one of the Kaggle playground prediction competitions, namely, dog breed identification. The supplied datasets include a series of dog pictures with corresponding breeds and a series of human pictures. The human pictures will serve as an initial gate (cascade classifiers)[1,2] for training the model to recognize whether the supplied picture is a dog. A pre-trained model VGG-16 (PyTorch) will then be implemented to detect dog breeds in the supplied picture.
The dog breed dataset includes the following components:

1. Train
   The training dataset contains 133 dog breeds with 30-70 pictures for each breed.
2. Test
   The test dataset contains the same 133 dog breeds with 6-10 pictures for each breed.

3. Validation

The validation dataset contains the same 133 dog breeds with 6-10 pictures for each breed.

The human dataset contains 5749 different pictures of human.

The majority of the images contains a single portrait of a dog of the corresponding breed. However, the dataset contains a small portion of images that include both dogs and humans, and other animals like cows and sheep. The challenge is how to successfully detect whether there is a dog in the image and how not to identify other animals or humans as dogs, or vice versa.

Solution Statement

The solution of this project will include a trained detector using CNN to detect dog breeds based on the supplied picture. The accuracy rate is expected to be above 90%. The model is expected to tell if the dog is pure breed or mixed breed based on feature scores. If the dog is a mixed breed, the model should show possibility scores of which breeds are mixed. The model is expected to tell human apart from dogs. As often used by the Kaggle competition, the model will be evaluated by the multiclass logarithm loss score.

Benchmark Model

The benchmark for the model can be referenced to the Kaggle leaderboard for dog breed identification competition. The target for this model is to reach a multiclass loss score less than 0.01, which is in the top 100 of the competition. The other benchmark will be 90% prediction accuracy.

Evaluation Metrics

This project will implement the same evaluation metrics as defined by the Kaggle dog breed identification competition, which is multi class log loss between the predicted probability and the observed target. The multi class log loss measures the performance of a classification model where the prediction input is a probability between 0 and 1. In this proposed dog breed classifier, the output will contain an array with 133 probability values of how likely the supplied image belongs to each of the 133 breeds. Therefore, the multi class log loss metric is perfect for this project. The goal is to minimize the multi class log loss, optimally to zero. The other metric is the accuracy of the prediction, which is naturally the most straightforward metric for the nature of this project.

Project Design

The project will follow the steps provided by the standard project template:

Step 0: Import Datasets.
Step 1: Develop component using Haar Cascade classifier to Detect Humans.
Step 2: Develop component using pre-trained VGG-16 model to Detect Dogs.
Step 3: Assemble components to a full detection algorithm.

Step 4: Test the completed algorithm.

The total dataset (include training, test, and validation) contains 133 breeds of dog with approximately ~8.3k of images.
Before constructing the CNN model for dog breed classification, the dataset is explored by testing the pre-trained model, i.e., cascade classifier and VGG-16 to detect whether the image contain humans or dogs. The images will be passed through the human and dog detectors before passing into the breed classifier. If no dogs are detected in the image, the image will still be passed into the breed classifier and the resultant most probable breed will be returned for fun context.

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.