Capstone Project Proposal: Dog Breed Classifier with CNNs

by Ali Mert Ertugrul - Nov 26, 2020

Domain Background

Image processing and computer vision are very active two areas of artificial intelligence (AI) especially with the recent advancement in deep learning. Dog breed identification or classification is a typical task in the image processing / computer vision domain. The purpose is to identify/classify the dog breed based on a given picture of a dog as the input. Such an application would be very useful for a large audience (e.g., children) who want to learn and investigate more about the dogs. This task has been widely studied in the literature by employing various techniques. For instance, Liu et al [1]. employed part localization technique to perform dog breed classification. In another study, Wang et al. [2] modeled the shape of dog breed as points on the Grassman manifold. Then, they performed dog breed classification on that manifold. More recently, Borwarnginn et al. [3] used Convolutional Neural Networks (CNNs) with transfer learning to classify dog breeds and compared the results with those of conventional methods. In this project, I also aim to build a dog breed classifier using CNNs and transfer learning, and evaluate this classifier on the dataset including 8351 images of 133 different breeds. Therefore, my purpose is to improve my machine learning skills (e.g., learning how to implement CNNs, using transfer learning) on the image processing domain by performing a well-known task.

Problem Statement

The problem within the scope of this project is basically to estimate the dog breed based on a given picture of a dog as the input. This is a multi-class classification problem. Therefore, the goal is to build a machine learning pipeline to process real-world, user-supplied images for this task. In general, this pipeline will take an image as the input and estimate the breed of the dog in the picture. More specifically, the pipeline will perform the following tasks based on a given image as the input:

1. Estimate the canine's breed based on the given image of a dog with a reasonable accuracy.
2. Estimate the most resembling canine's breed if the supplied image (input) is an image of a human.

Datasets and Inputs

For this project, two different datasets are provided by Udacity. Both datasets include images. The first dataset, called dog-image dataset, includes images of dogs where the second dataset, called human-image dataset, includes images of humans. The more detailed information is provided for the corresponding datasets as follows:

1. *Dog-image dataset:* It includes 8351 dog images 133 different breeds. The properties of the images are not consistent meaning that the image size, background and illumination may vary across the images. The numbers of the images for each breed are not uniformly

distributed. For some of the breeds the dataset contains more images. Furthermore, the dataset is divided into three sets, training, validation and test, where there are 6680, 835, and 836 images for training, validation and test sets, respectively. They will be used for training, validation and test in estimation of the breed in the images.

2. *Human-image dataset:* It includes 13233 human images from 5749 different people. For some people, there are more images whereas for some of them there are less images. Therefore, the dataset is not uniformly distributed with respect to the people. All images in the dataset have the fixed image size which is 250x250 although the background and illumination may vary across these images. They will be used to test the human face detected as well as estimating the resembling breed for the given human image.

Solution Statement

In this work, we will employ Convolutional Neural Networks (CNNs) to perform dog breed classification. Our solution will (1) predict the breed of the dog based on a given input image of a dog, (2) estimate the most resembling dog breed based on a given input image of a human. In our algorithm, we will follow the several steps as follows:

1. Detect whether a given input image contains a human. To do that, we will use one of the Haar feature-based cascade classifiers of OpenCV[1].
2. Detect whether a given input image contains a dog. For this purpose, we will use a pre-trained VGG16 model which is a CNN-based architecture trained on ImageNet [4].
3. Develop a dog breed classifier from the scratch using CNNs.
4. Develop a dog breed classifier using transfer learning.

Benchmark Model

I consider two different benchmark models for this task from the literature. It is important for benchmark model selection to obtain studies, methods that perform the same task on the same dataset. I select the published studies [1] and [2] which perform dog breed classification on the same dataset. They achieved 67%, 96.5% accuracy, respectively. I will compare the results of the proposed solution to those of the benchmark models.

Evaluation Metrics

Since this is a multi-class classification task, *Accuracy* will be the main evaluation metric in this work. It is also convenient to compare accuracy results with the benchmark models since they provide accuracy results as well. The accuracy is calculated as follows: the number of correctly predicted test samples is divided by the total number of test samples. In addition, I will evaluate the proposed solution with *micro- and macro-average Precision, Recall* and *F-score* since we have a class imbalance in the dataset.

[1] https://opencv.org/

Project Design

Our project consists of 7 steps to follow:

- *Step 0 - Download and import datasets:* Dog and human datasets will be downloaded and imported.
- *Step 1 - Detect humans:* For a given input image, an algorithm will be developed to detect if there exists a human in the image using a Haar feature-based cascade classifier of OpenCV.
- *Step 2 - Detect dogs:* For a given input image, an algorithm will be developed to detect if there exists a dog in the image. The algorithm will be developed using the pre-trained VGG16 model.
- *Step 3 - Create a CNN to classify dog breeds from the scratch:* A CNN model will be developed to classify the breed of the dog in the given input image. It will be implemented using PyTorch[2].
- *Step 4 - Create a CNN to classify dog breeds using transfer learning:* A CNN model will be developed using a pre-trained VGG16 or ResNet50 model as the basis. The model will be finetuned on the dog dataset in our work. PyTorch will be used as the framework to develop the model. The model will be evaluated on the test set.
- *Step 5 - Write you algorithm:* An algorithm will developed to perform the following tasks:
    - Detect if the given input image contains a dog. If so, return the predicted breed of the dog using the model developed in Step 4.
    - Detect if the given input image contains a human. If so, return the most resembling dog breed using the model developed in Step 4.
    - if neither is detected in the image, provide output that indicates an error.
- *Step 6 - Test your algorithm:* The algorithm developed in Step 5 will be tested on at least six images (at least two human images and at least two dog images).

References

[1] Liu, Jiongxin, et al. "Dog breed classification using part localization." European conference on computer vision. Springer, Berlin, Heidelberg, 2012.
[2] Wang, Xiaolong, et al. "Dog breed classification via landmarks." 2014 IEEE International Conference on Image Processing (ICIP). IEEE, 2014.
[3] Borwarnginn, Punyanuch, et al. "Breakthrough Conventional Based Approach for Dog Breed Classification Using CNN with Transfer Learning." 2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE). IEEE, 2019.
[4] Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." 2009 IEEE conference on computer vision and pattern recognition. IEEE, 2009.

[2] https://pytorch.org/