Machine Learning Engineer Nanodegree Capstone Proposal
Eze Chijioke Christian
June 18, 2020

Project Topic: Algorithm for a Dog Identification App

1. Introduction (Domain Background)

Object identification from images is one of the areas that has been witnessing great interest by researchers and industrial practitioners within the broad field of computer vision. This stems from its importance in video analysis and image understanding tasks (Zhao et al., 2019). It is worthy of note that the processes involved in identifying an object based on its image are several. Whereas these processes can be handled easily by the human brain, on the other hand, computers require more elaborate specifications and computations to be able to accurately identify an object from image file(s).

Dog is one of the most favorite pets of many households in the world and dog business is a huge business. One common problem encountered when dealing with huge dog data is the difficulty of determining the breed of a particular dog given several dog images. There are many attempts by researchers to solve this problem. For instance, (Liu et al., 2012) proposed part localization-based dog breed classification method that achieved 67% accuracy on a dataset they downloaded from Flickr, ImageNet, and Google (available at http://faceserv.cs.columbia.edu/DogData/). Similarly, (Rhodes, 2012) implemented a modified version of the method proposed by (Liu et al., 2012) on the same dataset with less success (55% accuracy). An improved result was obtained when they experimented with CNN on the dataset. Further, (Schuerger, 2013) implemented SVM with a linear kernel to classify dog breeds based on dog features extracted via scale-invariant feature transform (SIFT) descriptors and color histograms. In their work, facial keypoints of each dog image were extracted using a convolutional neural network (CNN). The accuracy of the method they proposed is approximately 50%. Other studies on dog breed exist, however, it can be generally concluded that the accuracies of most existing solutions are still far from impressive. Hence, in this project, we aim to leverage on the existing solutions to develop a more accurate solution to the dog breed classification problem. Finally, we created an algorithm that, given an image, it will be able to check if its dog, human, or other object and generate a response accordingly.

2. Problem Statement

There has been a growing interest in the design of web or mobile app for use in identifying dog breeds from real-world, user-supplied images. While such an app has huge potential benefits in terms of fast identification and improved accuracy, it is also not an easy task to do. This is because the data may be contaminated, that is, the dog image data may contain images of other species. An instance of those other species is a human being. Thus, it becomes necessary to design a solution that would perform well even in the presence of human faces or other objects within the supplied image data. This project aims to address this problem through the use of machine learning techniques. This can be realized by designing models to detect if the supplied image is a human face, dog, or other objects. Further, given an image, an algorithm developed as part of this project would be able to detect the type of dog breed the supplied image belongs to: if the image is a human image, a statement is printed indicating that a human face is detected as well as stating a particular dog breed the human image resembles.

3. Datasets and Inputs

The datasets to be used in this project are the provided datasets. More details about the datasets are as follows:
i. Dog dataset (https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/dogImages.zip):
This will be used to train and test the dog detector model that will be developed in the project.
This dataset is necessary in developing the dog detector app because the proposed models need to be trained with sample dog images to be able to generalize and detect breeds from previously unseen dog images. Information about the dataset (**dogImages)** as obtained from data analysis (attached to the project folder) are as follows:

- **train** → no. of training images: 6680 (80%)
- **valid** → no. of validation images: 835 (10%)
- **test** → no. of testing images: 836 (10%)
- **number of classes:** → 133

A further, analysis of the **train** shows the following statistics.

- **Class of dog distributions in the training dataset**
  - Mean of the dog beed class counts in the training dataset is: 50.23
  - (min dog beed class counts, max dog beed class counts): (26, 77)
  - The STD of the dog beed class counts in the training dataset is: 11.82
- **Image Size distributions in the training dataset**
  - (avg width, avg height): (571, 532)
  - (min width, max width): (112, 4278)
  - (min height, max height): (113, 4003)
  - (STD of widths, STD of heights): (397, 342)

N.B: A detailed analysis of the datasets with several plots can be seen in the included data analysis files (data_analysis.pdf and data_analysis.ipynb).

ii. Human dataset (https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/lfw.zip):
This dataset will be used to test a human face detector model that will be used to test the human face detector part of the solution. Because the proposed solution needs to be able to detect if a supplied image is human face instead of dog image, this dataset will be used to test a pre-trained model to be used in performing this task.

4. Solution Statement

This study will lead to the development of a human detector based on OpenCV's pre-trained human face detector model. Afterward, we will use VGG-16 model, together with weights that have been trained on ImageNet (http://www.image-net.org/). ImageNet contains over 10 million URLs, each linking to an image containing an object from one of 1000 different objects (https://gist.github.com/yrevar/942d3a0ac09ec9e5eb3a) to detect dogs. The accuracy of this model on the provided dataset will be noted. A CNN model will then be developed from scratch and tested based on the dog dataset to detect dog from an image and the accuracy score of the model will be noted too. Further, Finally, a CNN to classify dog breeds based on transfer learning approach will be developed.
The best dog breed predictor model from above will be used in the final solution. Thus, at the end of this project, an algorithm that accepts a file path to an image, determines if the image contains a human, dog, or neither will be developed. Particularly,

- if a **dog** is detected in the supplied image, the algorithm returns the predicted breed of the dog.
- if a **human** is detected in the supplied image, it returns a dog breed the person resembles.
- if **neither** is detected in the image, the algorithm generates an output indicating an error.

5. Benchmark Model

The accuracy of the final dog detector model will be compared with other existing results reported by other researchers. Table 1 shows a few of these results. Performance of the proposed method will be compared with these existing works and we intend to significantly surpass them all.

Table 1: Results of previous dog detector models.

Work	Method	Accuracy Score (%)
(Liu et al., 2012)	Part localization approach	67.00
(Rhodes, 2012)	Modified part localization approach	55.00
(Schuerger, 2013)	SVM, SIFT, and CNN	50.00

6. Evaluation Metrics

Accuracy score will be the metric to be used to evaluate the performance of the developed solution. Accuracy by definition can be derived a s follows:
Let
TP = true positive
TN = true negative
FP = false positive
FN = false negative

Thus, accuracy can be evaluated as

Accuracy = (TP + TN) / (TP + TN +FP + FN) .

However, in the case of the project, the accuracy of the model can be evaluated by determining the to number of correctly classified dog images and dividing it with the total number of all tested images.

Justification for choosing accuracy: This is because the training dataset is fairly balanced (with STD of the dog beed class counts in the training dataset of 11.82) as reported above under Datasets and Inputs,.

7. Project Design

The workflow outlined in Figure 1 above will be followed to realize this project. First, the datasets will be loaded in a jupyter notebook and a few of the images will be visualized to understand the nature of the data. Second, a program to detect human face referred to as human detector will be developed using OpenCV library. This program will be tested with some samples of the human image dataset. Third, a program using a pre-trained VGG-16 model will be developed to detect dogs. Fourth, a CNN model will be developed to detect dogs and tested too. Further, a CNN developed using transfer learning approach will be developed. The best of the dog detector models will be selected based on its test performance. Finally, an algorithm utilizing the selected best dog detector model and human detector program will be developed to solve the dog detection problem. Some dog and human images will be used to test the final solution to ensure that it is working properly.

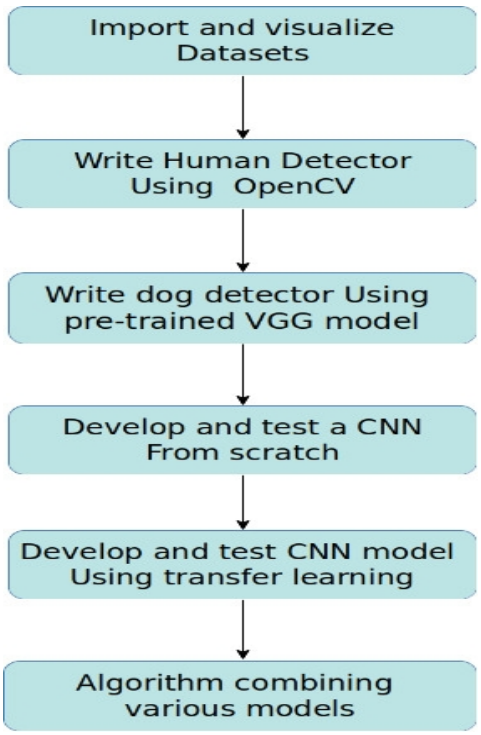

Figure 1: Dog Breed Classification Workflow

References

Liu, J., Kanazawa, A., Jacobs, D., & Belhumeur, P. (2012). Dog breed classification using part localization. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7572 LNCS*(PART 1), 172–185. https://doi.org/10.1007/978-3-642-33718-5_13

Rhodes, D. (2012). Automatic Dog Breed Identification. *Automatic Dog Breed Identification*, *1*(1), 1–6. http://cs231n.stanford.edu/reports/2015/pdfs/automatic-dog-breed.pdf

Schuerger, R. J. (2013). Dog breed identification. *Journal of the American Veterinary Medical Association*, *242*(1). https://doi.org/10.2460/javma.242.1.27

Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object Detection with Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(11), 3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865