Machine Learning Engineer Nanodegree Capstone Project Proposal

Proposal

To design and implement a CNN-based Deep Neural Network (DNN) model to perform the task of classifying different breeds of dog based on an input image provided by user. The performance of this model will also be evaluated using the performance obtained from an existing CNN-based DNN model as a benchmark.

Domain Background

The domain in which this project falls into is Computer Vision and Object Classification. Since the advent of modern computers built on semiconductor devices, computer scientists have been attempting to apply mathematical algorithms to be executed automatically by computers to address a variety of tasks that seems trivial to human such as lesser complex problems such as line and edge detections in images to more complex problems such as either single or multiple object classifications or detections in images.

For performing complex tasks such as object classification or recognition from the perspective of using computers, a very important aspect is the selection of suitable features that will provide sufficient information to distinguish/discriminate one object class from another. As such, great emphasis has been placed on Feature Engineering and Extraction in both spatial and frequency domain such as Corner features, HAAR-like features, Sift Invariant Feature Transform (SIFT) to Speeded Up Robust Features (SURF). However, with the introduction of Convolutional Neural Networks (CNN) such as AlexNet in 2012, the importance of feature engineering for obtaining good object classification results in imagery has decreased significantly over the years. This is due to the fact that a pretrained deep CNN (trained using tens of thousands or hundreds of thousands of images) will be able to extract much more robust and complex features that can provide object classification or recognition performance that are even superior to a standard human performance.

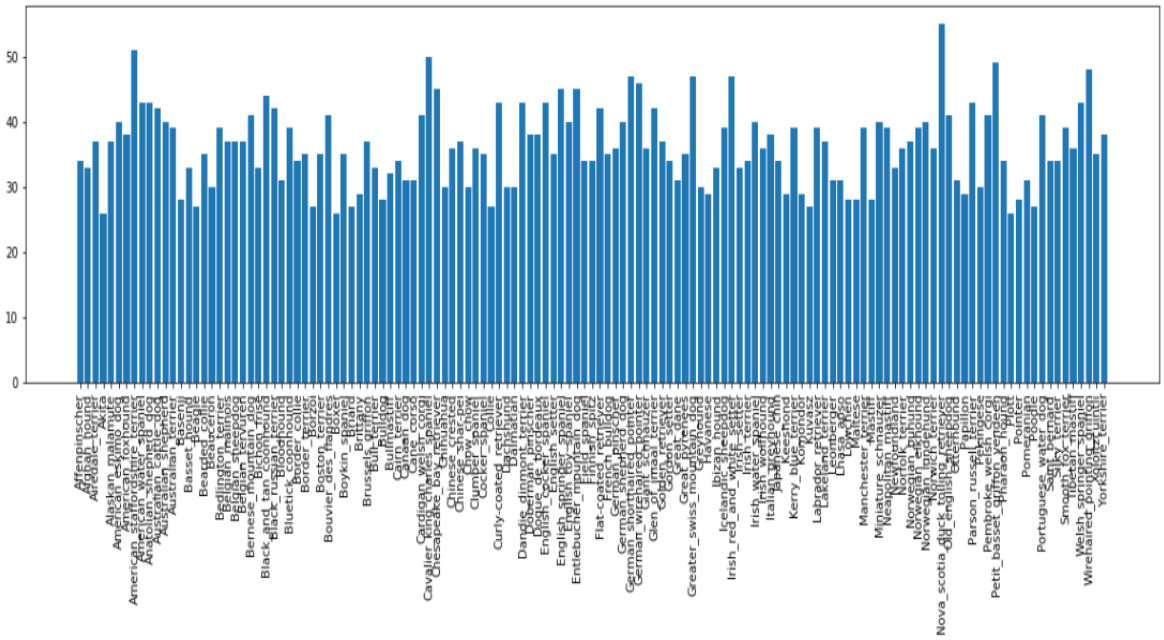
Problem Statement

In this project, the objective is to demonstrate the capability to design and implement a CNN-based Machine Learning Classification pipeline for a multi-category classification model to provide a classification result based on the highest prediction probability for each user-supplied image. To be more specific, when provided with an image of a canine, this Machine Learning Classification pipeline will provide a prediction of the **canine's breed** such as an Australian Terrier or Beagle based on the breed with the highest probability computed from the output of the CNN model. However, if the user-supplied image belongs to that of a human, the Machine Learning Classification pipeline will then attempt to associate this human input image with the closest resembling dog breed.

Datasets and Inputs

In this project, the data to be used for training, validating and testing the implemented CNN-based Dog Breed Classifier as well as for obtaining the performance of the bench marked model are provided by Udacity. These data fall into two main groupings, namely images of different breeds of dogs grouped under the dog grouping and images of different human faces grouped under the human grouping.

By performing some exploratory data analysis, under the dog grouping, the dataset provided is located within the **dogImages** folder containing **8351** dog images and belonging to **133** dog breed categories. Within this folder, the dog images are allocated into 3 sub folders named **test**, **train** and **valid** with 836 images under test sub folder, 6680 images under train sub folder and 835 Images under valid sub folder. Also, a snapshot of the distribution of the breed of dog images under the train sub folder is shown below:



Next, under the human grouping, the dataset provided is located within the **ifw** folder containing **13233** human images under **5750** sub folders with each sub folder belonging to a person. Also, each human image within the human grouping has a fixed size of 250 x 250 pixels.

Finally, these data are available under the following weblinks:

Dog data available at: https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/dogImages.zip
Human data available at: http://vis-www.cs.umass.edu/lfw/lfw.tgz

Solution Statement

In this project, two approaches are considered implementing the Machine Learning Classifier for Dog Breeds, namely traditional Machine Learning Algorithms or Convolutional Neural Network models.

When referring to traditional Machine Learning Algorithms, they can be less sophisticated algorithms such as Logistic Regression or Support Vector Machines to more sophisticated algorithms such as Naïve Bayes, Random Forest, Decision Trees or XGBboost. Although these algorithms can provide good performance for problems of varying complexity, a feature selection process will need to be applied to identify the features to be used for these algorithms. As such, the solution to be adopted will be based on CNN models that will incorporate the feature selection process as part of its training process.

Benchmark Model

In this project, the benchmark model to be used for comparison will be the pretrained **VGG-16** model that has been trained using the ImageNet database which has millions of images for the 1000 categories of objects and humans. Thus, the performance obtained from applying this model to the Dog Breed Classification will be a good objective comparison to the proposed solution. Alternatively, a second pretrained model such as the original **AlexNet** model can also be deployed to quantify its performance against that of the proposed solution. This suggestion will show the progression of Artificial Intelligence capability of CNN-based models within the span of a decade from 2012 to 2020.

Evaluation Metrics

In a standard classification problem, the default evaluation matric that is used to quantify the performance of both the benchmark model and the solution model will be the **Accuracy Score** in % in correctly predicting the correct breed of dog from the test data provided. For instance, if the test data set contains a total of 150 dog images and 90 of these dog images are assigned with the correct dog breed, then the Accuracy Score will be 90/150 * 100% = 60%.

However, in this project, from the results obtained of the exploratory data analysis for the train sub folder as shown previously, it can be seen that it is not a balanced dataset as there are dog categories with more dogs than others. As such, instead of the Accuracy score, other metrics such as the precision **P** and the recall **R** will be used. Also, **P** is defined as the number of correct positive results divided by the number of all positive results returned by the classifier and **R** is defined as the number of correct positive results divided by the number of all relevant samples (all samples that should have been identified as positive) (cite from https://en.wikipedia.org/wiki/F1_score).

In addition, a **Confusion Matrix** may also be provided to illustrate the distribution of correct classification results versus incorrect classification results so that more insights are provided for the classification results. (cite from https://en.wikipedia.org/wiki/Confusion_matrix)

Project Design

For this project, the intended workflow will be as follows:

1) Data Analysis: Perform some Exploratory Data Analysis on the dog image dataset so as to understand the size of the images and images available under each breed category etc.

2) Perform the steps in the order as listed in the Jupyter Notebook provided by Udacity for this project:
    i. Evaluation of Human Detector using existing OpenCV pretrained models and Evaluation of Dog Detector using existing pretrained CNN-based models
    ii. Creating Data Loaders that includes necessary preprocessing process for the training, validation, and test datasets of dog images to be used for training the Custom designed CNN-based Dog Breed Classifier
    iii. Describe the details in choosing the various layers for building up the Custom-designed CNN to classify Dog Breeds
    iv. Train the Custom CNN and then perform testing of this trained Custom CNN on the test dataset of dog images to achieve an Accuracy Score of at least 10%
    v. Use a pretrained CNN model to perform Transfer Learning to obtain an Accuracy Score of at least 60% for classifying the test dataset of dog images
    vi. Develop the Function or App that will perform the following:
        - Accepts a file path to an image and first determines whether the image contains a human, dog, or neither
        - If a dog is detected in the image, return the predicted breed
        - If a human is detected in the image, return the resembling dog breed
        - If neither is detected in the image, provide output that indicates an error
    vii. Test out this function or app on at least 6 test images with two of these images to be dog images and two images to be of human
    viii. Answer all the questions that are listed in the Jupyter Notebook

3) Prepare the Final Report that fulfills all the required Rubrics for the Capstone Project for submission.