Proposal already reviewed at:
https://review.udacity.com/#!/reviews/2179839

Using Deep Learning techniques for Dog Breed Classification – A capstone project proposal

DOMAIN BACKGROUND

The goal of this project is to make a classifier for Dog Breeds. The motivation behind this choice came from my curiosity to work in the field of Computer Vision, and how that can greatly affect people. The task in hand is the one of Image Classification, and for that, Convolutional Neural Networks, or, as commonly said, CNNs, are usually used. Nowadays all new image classification networks architectures are based on this type of network. [2]

Convolutional Neural Networks are a kind deep learning neural networks that implements the so-called Convolutional Layers. Theses Layers, combined with a pooling layer, are responsible for dimensionality reduction and feature extraction [1][Figure 1]. The idea is so that by a set of trainable filters the network fits to dataset and can infer visual features [1]. Example: A CNN that is used to detect birds will be able to make filter that highlight beaks or feathers [Figure2].

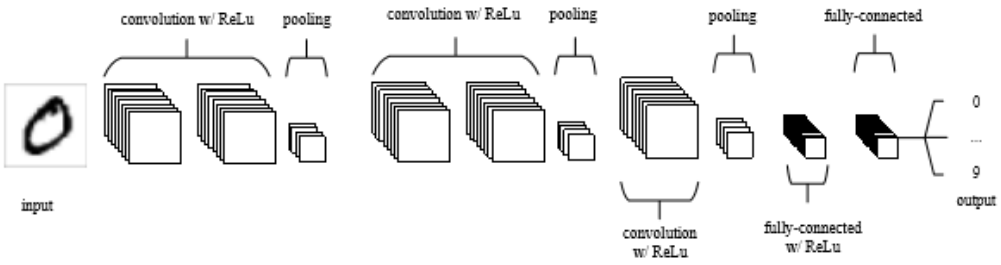

Figure 1 - Basic CNN architecture [1];

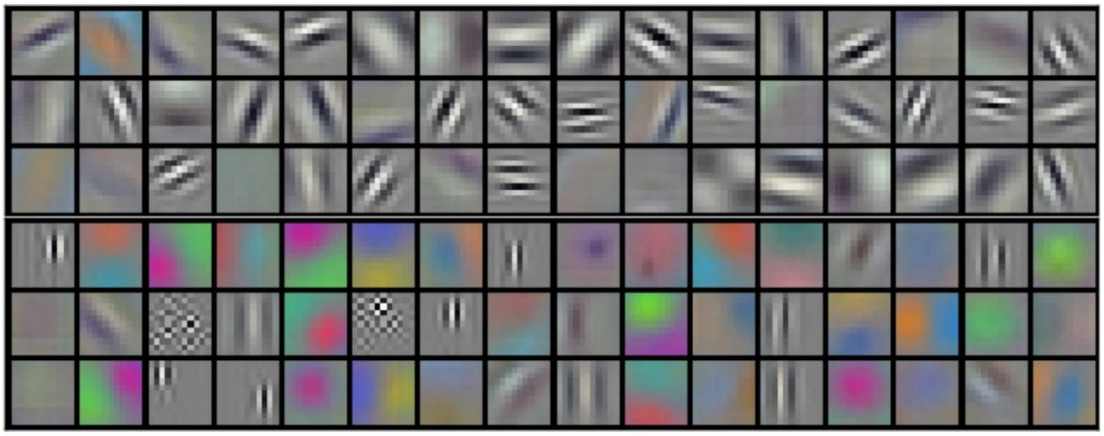

Figure 2 - Filters of a Layer. Notice all the forms, textures and colors these filters get.

THE PROBLEM AT HAND

It's expected that an image classifier is constructed using CNNs based architectures. But why construct such classifier? Have you ever wondered what dog breed is that one of your dog's friend or if the dog that you rescued have a pedigree of sorts prior to having to go to a veterinarian? And that's the presumption that I will take on this project.

THE DATASET

The Dataset that will be worked on consists of 8351 images of dogs that makes part of 133 different breeds as can be seen in https://www.kaggle.com/c/dog-breed-identification/data. It's presumed that the Dataset is already curated, so it should hold no problems with skewed data. And besides that, the dataset is already separated in train/test/validation folders. These images will be the input for the classification model that will be constructed.

One problem that can be present in the dataset but might be easily resolved is that one of the quantity of data points for each class in the dataset, but a data augmentation should be enough to address this problem and that will be made to achieve better classification results.

EVALUATION METRICS

It's expected that the model achieves at least 60% accuracy on this dataset, what that means it's that it's working better than a flip coin. I would expect a success characteristic that this would hold true to every class aswell, what that means is that a recall larger than 50% is desired aswell.

The best model classification of this dataset on Kaggle, by this day, is [https://www.kaggle.com/phylake1337/0-18-loss-simple-feature-extractors] uses transfer learning technique in a ResNetV2 deep learning architecture and achieves a training loss of 0.1705, training accuracy of 0.9461, and validation loss and accuracy of 0.1966 and 0.9404 respectively.

PROJECT WORKFLOW

The goal is to use the dataset as it is to achieve the highest metrics possible. This will be achieved by using transfer learning from a VGG16 on the constructed model and to thoroughly test and look out for the metrics. Besides that, the model prediction will be available as a function, at first, so that it could be evolved later to an endpoint.

REFERENCES

[1] O'Shea, Keiron & Nash, Ryan. (2015). An Introduction to Convolutional Neural Networks. ArXiv e-prints.

[2] VGG16 – Convolutional Network for Classification and Detection. https://neurohive.io/en/popular-networks/vgg16/