

**Capstone Project Proposal:**  
**CNN Dog Breed Classifier**  
**Javier Martinez**

Domain Background:

The field of research where this project belong is computer vision. Computer vision could be defined as the process whereby a machine could get information about some image and transform this data into the something that is intelligible to the human. According to Wikipedia computer vision *“is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images. It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding”*.

This task is very difficult, because requires and high level of abstraction that is created by the human and because of that is almost impossible to build a system than can resolve the problem with traditional programming.

It is commonly accepted that the father of Computer Vision is Larry Roberts, who discussed the possibilities of extracting 3D geometrical information from 2D perspective views of blocks (polyhedral) (Aloimonos, 1992). Since then, there have had a lot of progress, that includes change both algorithms and hardware

In recent time, the main developments could be resumed in two names: Convolutional Neural Network and GPUs. A convolution neural network is a special deep learning structure that allow create new features, like edges. This bring a lot of the abstraction and allow to the model adjust the parameters that are needed to create a good model. However, that requires a lot of computational power that means hours or may be moths of training. Fortunately, many of these operations can be done in parallel and with the use of GPUs<sup>1</sup> it's possible to train a CNN in less time. (NG, 2019)

The develop of this field has allowed to improve different systems, like systems that can find images related to a search, several image recognition systems that are used to identify people, systems that can identify disease in x rays images, etc.

---

<sup>1</sup> Google also developed a TPU, a special hardware that is very fast to process deep learning neural networks. You could use also for free a little instance in google colab: <https://colab.research.google.com>

My main interest in this topic is continue developing skill in the processing of images. I think there a lot o possibility of build system that could help in the identification of fraud, for example, in the recognition of people that commit presential fraud.

### Problem Statement:

As were stated in the Udacity page the problem will be “to learn how to build a pipeline to process real-world, user-supplied images. Given an image of a dog, your algorithm will identify an estimate of the canine’s breed. If supplied an image of a human, the code will identify the resembling dog breed.”

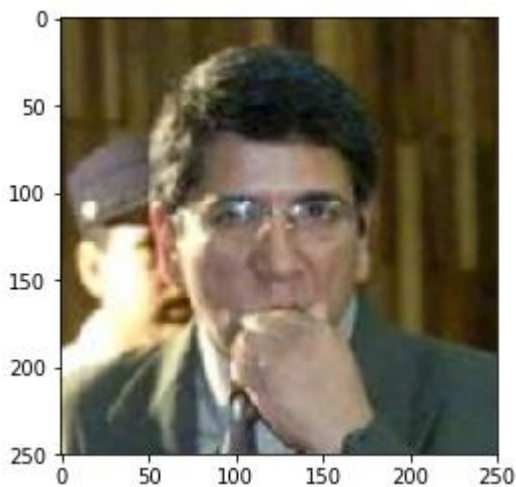
### Datasets and Inputs:

The dataset was share by Udacity. The dataset includes 13233 images of humans and 8351 images of dogs.

Every image of humans is of size 250 x 250 x 3

The images of dogs don’t have a constant size.

Below I share two examples of the pictures (index -45)



This dataset will be used first to identify if the picture is a dog or human, and after that, establish what are the breed of dog that is more related with this picture.

### Solution Statement:

We will need to solve two problems:

- First, we will find if the image is a human or a dog with two pretrained models:
  - Haar feature-based cascade classifier to find humans
  - VGG-16 Model to find dogs
- Second, we will use a CNN that will identify which breed is more related with the image.

### Benchmark Model:

After a search in internet I find the following table (Hsu) that resume a benchmark for the problem:

Method [citation]	Top - 1 Accuracy (%)
SIFT + Gaussian Kernel [1]	22%
Unsupervised Learning Template [7]	38%
Gnostic Fields [10]	47%
Selective Pooling Vectors [5]	52%

### Evaluation Metrics

As the notebook from Udacity stated, we will measure all the models with the accuracy metric.

## ***References***

- Aloimonos, Y. (1992). *Special Issue on Purposive and Qualitive Active Vision, CVGIP B*. Image Understanding.
- Hsu, D. (n.d.). Using Convolutional Neural Networks to Classify Dog Breeds. *Stanford University*. Retrieved from [http://cs231n.stanford.edu/reports/2015/pdfs/fcdh\\_FinalReport.pdf](http://cs231n.stanford.edu/reports/2015/pdfs/fcdh_FinalReport.pdf)
- NG, A. (2019). Convolutional Neural Networks in TensorFlow Cousera Course. Retrieved from <https://www.coursera.org/learn/convolutional-neural-networks>