

# Capstone Proposal

## Dog Breed Classification

**Yuhou Zhou**

zyh.germany@gmail.com

Machine Learning Engineer Nanodegree, Udacity

### 1 Domain background:

Image classification is a long standing problem in computer vision. In recent years, convolutional neural network (CNN) becomes the most popular algorithm to tackle this problems, because its high accuracy and the increasing of computer power enabling to build deeper neural networks. In 2012, AlexNet marked the new era of CNN. Its performance greatly surpassed other manually featured neural networks in many benchmarks. Afterwards, models, such as ResNet, MobileNet, Inception, all drew great attention because of their exceptional performance.

### 2 Problem statement:

This project aims to build a dog breed detector. When the input image is a dog, the detector returns the breed of the dog; when the input image is a human, the return should be human and give a dog breed which the human mostly looks like. If neither the cases, the detector should give a result indicating that neither of the are cases detected.

### 3 Datasets and inputs:

The datasets are provided by Udacity. The dog image dataset includes 8351 images and is divided into training set (6680 images), validation set (835 images), and test set (836 images). All the directories are further divided into 133 sub-directories, corresponding to different dog breeds. The human image dataset contains 13233 human images.

### 4 Solution statement:

There are many existing methods to solve image classification problems. We can try handcrafted models, pre-trained models, building a CNN from scratch, using transfer learning to fine tune many pre-trained models.

### 5 Benchmark model:

The CNN model built from scratch must reach 10 percent accuracy rate. Though the requirement seems low, considering the complex input space, the problem is not trivial.

The transfer learning based CNN have to reach 60% accuracy rate.

### 6 Evaluation metrics:

Cross Entropy Loss is used by the optimizer to learn the feature, because the model output the classification by the probability of every class.

Top-1 accuracy rate is used to evaluate the model, because our app is to output the most likely breed to user. It is important the the correct class has the highest probability.

### 7 Project design:

- Importing image datasets and inspect images. Dividing data into train, validation, and test sets.
- Preprocessing images for models.

- Detecting human face by OpenCV cascade classifier.
- Detecting dog by pretrained VGG 16.
- Building CNN from scratch to detect dog breeds. The accuracy should be at least 10 percent.
- Utilizing transfer learning to build a CNN based on ResNet. The accuracy should be at least 60 percent.
- Writing a function to detect human and dog breeds, wrapping up the dog detector and the human detector.