

CPE480-05  
Artificial Intelligence

Project Proposal  
Dog Breed Identification

**Group Members**

Alvin Chui  
Robert Hensley  
Justin Cho  
Toan Pham

**Instructor**

Hisham Assal

**Date**

October 7, 2019

## Overview

The purpose of this project is to develop an AI model that is capable of analyzing and identifying a dog by its breed using a captured photo or photo library. Using machine learning, we hope to greatly reduce the work required in creating an agent that detects dog breeds. Instead of manually inputting rules to find features that distinguish dog breeds, we will use a Convolutional Neural Network to find distinct features for us faster and with more accuracy.

## Team Members & Division of Labor:

Alvin Chui

- Set up system architecture for the project
  - Select platforms that can successfully run our models (best GPUs to build the models efficiently)
  - Select the best libraries for our project
  - Select image datasets that are best for our three models
  - Feature engineering of this data for our models

Robert Hensley

- Training model for identify if an image is of a dog or not a dog
  - Set up images from ImageNet or another image database to get non-dog images for negative testing
  - Parse through this data to create a reasonably sized dataset to create a strong dog detection model

Justin Cho

- Train a model that detects certain breed groups (~10 labels)
  - Go through the Kaggle dataset and create ~10 breed groups (such as hounds, terriers) and label them programmatically
  - Train a model and test its validity on unlabeled images
  - Change the amount of groups depending on the success or failure of the model

Toan Pham

- Train a model that detects specific dog breeds (120 labels)
  - This model will use the 120 labels given in the kaggle dataset
  - Train a model and test its validity on unlabeled images

## Expected Outcome

- Our final model should be able to identify whether an image contains a dog with an accuracy of at least 90%
- Model should be able to distinguish breed with an accuracy of at least 85% accuracy for each breed we are training on

- The final number of breeds we have good results for should be roughly 20 but more depending on our progress
- Training results will be output as a .csv file

## Test Plan

Our models will be tested against a testing data set to validate the accuracy of the system. An accuracy of over 90% on dog-breed-identification and 85% on object-identification is sufficient to be considered as a success. We will perform two tests of both of our model: object-identification and dog-breed-identification.

## Main Components of the System

Our system will contain three models: one for dog classification (is it a dog or not), dog breed-group classification and dog-breed classification.

These will be the python libraries we will use for our project:

- Tensor Flow
- Keras
- Jupyter
- SciPy
- Pandas
- Numpy
- OpenCV

These will be the algorithms we will use to create our model:

- Neural Network for Classifying dog images
  - In particular, we will probably utilize a Convolutional Neural Network (CNN) because it is a model that is often used for accurately classifying image data
- Matrix Operations for Data Augmentation of Tensors (our image data)
- Data Visualization Methods for analyzing our Model's Accuracy

## Data Sources

- For our dog-breed-identification model, we will be using the dogs dataset provided on Kaggle.com. Link: <https://www.kaggle.com/c/dog-breed-identification/data>
- For our object-identification model, we will be using the dataset provided on image-net.org