

# Dog Breed Identification

Minxi Duan (duan23m@mtholyoke.edu)  
Tamás Bukits (bukits98@gmail.com)

## 1. Introduction

The main goal of this project is to find a solution to classify dog breeds with a deep learning approach. Given pictures of different dogs, the solution should identify which breeds the dogs are. Currently, 120 different breeds are supported in our training set, but later, the range of supported breeds could be expanded based on needs. In order to have a better user experience, we develop an android application for detecting dog breeds using the above solution we have. The application source code can be seen in the Github link.

## 2. Previous solution

There is a similar task - “Dog vs Cat identification” which is a standard computer vision dataset classification. This project only allow to identify general dog and cat, based on this project, we want to go deeper with neural convolutional networks. As a result, we choose to make a more detailed classification.

## 3. Dataset

Our datasets are downloaded from <https://www.kaggle.com/jessicali9530/stanford-dogs-dataset>. This Kaggle datasets contains 120 directories which indicates 120 breeds of dogs and a total number of 20580 photos. These datasets are kind of hard to predict as certain breeds of dogs have nearly identical characteristics, or differ in color and age.

We also upload the dataset to Github link. We split each directories into 3 folders which are the test, train and valid datasets with a 1:7:2. When splitting the data we tend not to mix the dog breed but keep them in their original directories.

## 4. Proposed method and Training datasets

We use a Convolutional Neural Network (CNN) specifically Inception v3 for this classification problem. Inception v3 is “a convolutional neural network for assisting in image analysis and object detection, and got its start as a module for Googlenet”. We try to train the model three convolutional layers, two dense layers and one layer for classification but failed with only a 10 percent accuracy, primarily because we only have around 20,000 photos in dataset which does not contains enough photo for this model to reach to a high accuracy.

As a result, we switch to transfer learning, which does not require that much labelled datasets to train complicated models. In the transfer learning model, add one more dense layer with 1024 neurons and one more layer for classification. We use the GlobalAveragePooling2D to get the average number of convolution values added into the model and add the dense layer with activation function “relu” and kernel\_initializer=‘he\_normal’. By this, we reach an accuracy of 0.8411

For training the model, we settled on a batch size of 20 and our first training validation accuracy achieved 82.10% from 57.67% after three epochs. With further iterating the process of the training and we managed to achieved an accuracy of 84.11% from 73.20%.

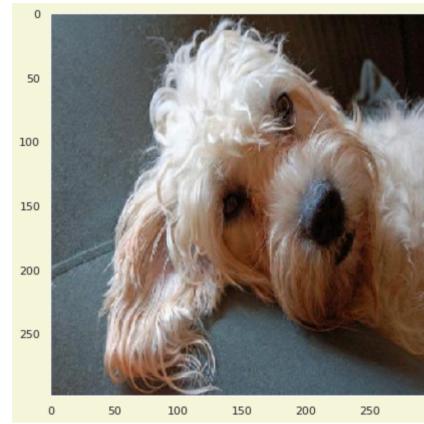
## 5. Evaluation method

When evaluate the model, we finally get a test accuracy of 86.20%. Among all the prediction we have, we got prediction True - 1821 photos(0.8622159090909091) and prediction False - 291 photos.

For each photo, the model will find the top three prediction and finally choose the prediction with higher similarity the algorithm detected.

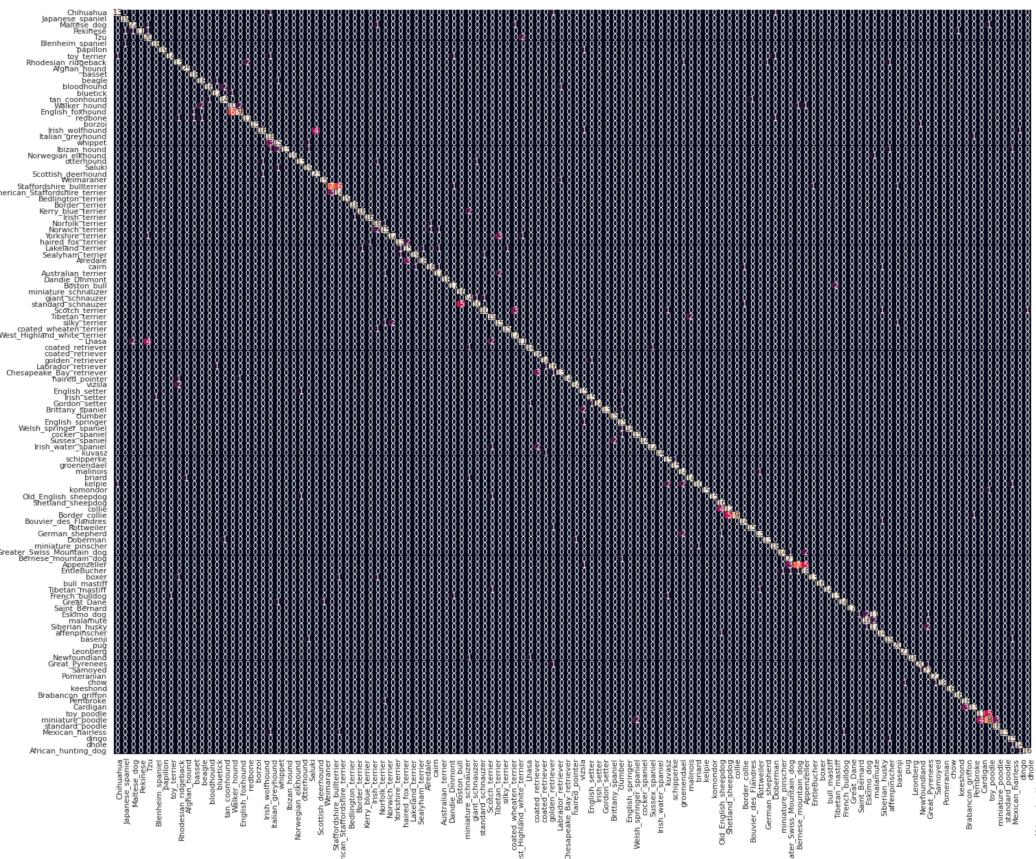


Top three predictions: Lhasa (95.40%), Tibetan\_terrrier (2.21%), Tzu (1.35%)  
Prediction: Lhasa  
Ground truth: Lhasa

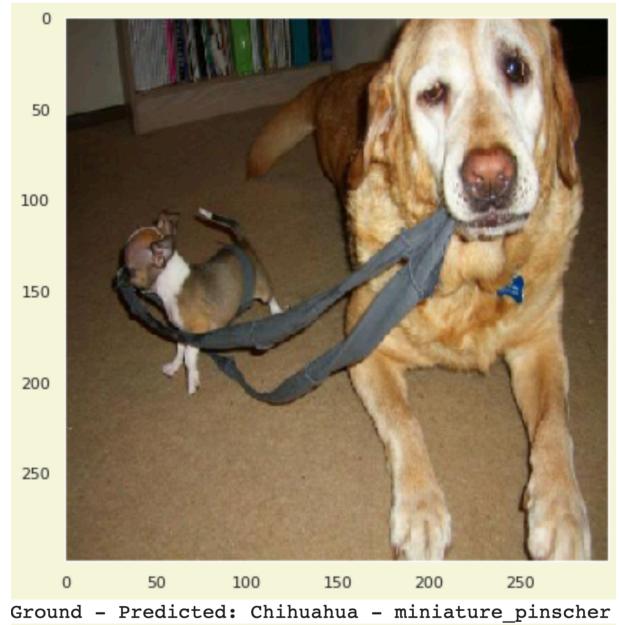
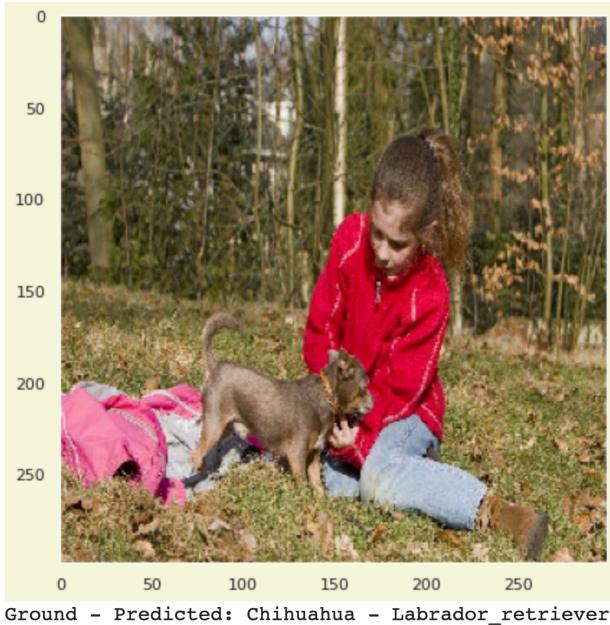


Top three predictions: Dandie\_Dinmont (41.33%), Sealyham\_terrier (35.22%), miniature\_schnauzer (13.38%)  
Prediction: Dandie\_Dinmont  
Ground truth: Dandie\_Dinmont

In order to figure out the result of our model, we made a confusion matrix. In the confusion matrix, we can see that the main diagonal where the prediction matches the actually dog breed were generally accurate with some weak spots. The misclassification range was from 1 to 11 and average around 3. (Larger photo find in Colab, link on Github)



To find out why these photo were predicted inaccurately, we summarize all the misclassification, and labeled the ground truth and the prediction. Here are some sample outputs:



## 6. Results

From the misclassification assemble we made, we classified them into three different condition.

First condition, the prediction is almost closed to the actual results with since the prediction have nearly identical characteristics, or just differ in color and age with the actual dog breed. Such as the reality is Maltese and prediction is the white Miniature Poodle.

Second condition, there is a large discrepancy between prediction and reality. Such as the reality is Chihuahua and prediction is Labrador Retriever.

Third condition, there are multiple dogs in the photo and our model did not detect the target one but the other dog which is different from the label.

The confusion of the first condition are obvious, the higher similarity between the two breeds, the higher possibility of making wrong classification. To improve the accuracy, we probably need a really large dataset to train the model. The confusion generate from the second condition are caused by different reason. These photos generally have lower pixel and a lot of interference factor such as a lot of people around them, the dog is stand near a balustrade of similar color, etc. In the third condition, since we don't choose and train the model with multiple target, so obviously, the random detected target might not be the labeled one. We could probably improve this in the future.

## 7. Discussion

Our project find a solution to classify dog breeds and reach a 86.22% test accuracy. In this report, we explained our project process and explain the possible explanation of false prediction.

## Acknowledgements

1. Aditya Khosla, Nityananda Jayadevaprakash, Bangpeng Yao and Li Fei-Fei. Novel dataset for Fine-Grained Image Categorization. First Workshop on Fine-Grained Visual Categorization (FGVC), IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011. [pdf] [poster] [BibTex]
2. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database. IEEE Computer Vision and Pattern Recognition (CVPR), 2009. [pdf] [BibTex]
3. <https://www.kaggle.com/jessicali9530/stanford-dogs-dataset> Kaggle datasets
4. <https://paperswithcode.com/method/inception-v3> Papers with code definition