Convolutional Neural Networks (CNN) project - Dog Breed Identification

Domain Background

The first case of neural networks was in 1943, when neurophysiologist Warren McCulloch and mathematician Walter Pitts wrote a paper about neurons, and how they work. They decided to create a model of this using an electrical circuit, and therefore the neural network was born.
In 1950, Alan Turing created the world-famous Turing Test. This test is fairly simple - for a computer to pass, it has to be able to convince a human that it is a human and not a computer.
1952 saw the first computer program which could learn as it ran. It was a game which played checkers, created by Arthur Samuel.
Frank Rosenblatt designed the first artificial neural network in 1958, called Perceptron. The main goal of this was pattern and shape recognition.

Much later mathematicians, scientists and engineers realised that GPUs could be used to improve the speed of computations used in their discipline, due to the use of vectors. This led to the realisation that GPUs would make neural networks, a very old idea, leaps and bounds more practical. This led to GPU companies particularly Nvidia benefitting hugely from the “machine learning revolution”.
From then to now machine learning has evolved to artificial intelligence and here we are detecting human emotions, designing robots and self-driving cars. That is amazing!

For this capstone project I am interested in Deep Learning. Deep Learning Deep learning is a type of machine learning, which is a subset of artificial intelligence.

Machine learning is about computers being able to think and act with less human intervention; deep learning is about computers learning to think using structures modeled on the human brain.
Machine learning requires less computing power; deep learning typically needs less ongoing human intervention.
Deep learning can analyze images, videos, and unstructured data in ways machine learning can’t easily do.
Every industry will have career paths that involve machine and deep learning.

Source:
doc.ic.ac.uk/~jce317/history-machine-learning.html
https://flatironschool.com/blog/deep-learning-vs-machine-learning

Problem Statement

The problem we are going to solve is to train a deep learning model to identify dog breed. This is an image classification problem.
This can have many applications that can be both helpful and fun especially for dog lovers, industries and companies that offer products and services for dog owners. It also has a fun part where when the model is given a human image it mentions dog breed that the human resembles to.

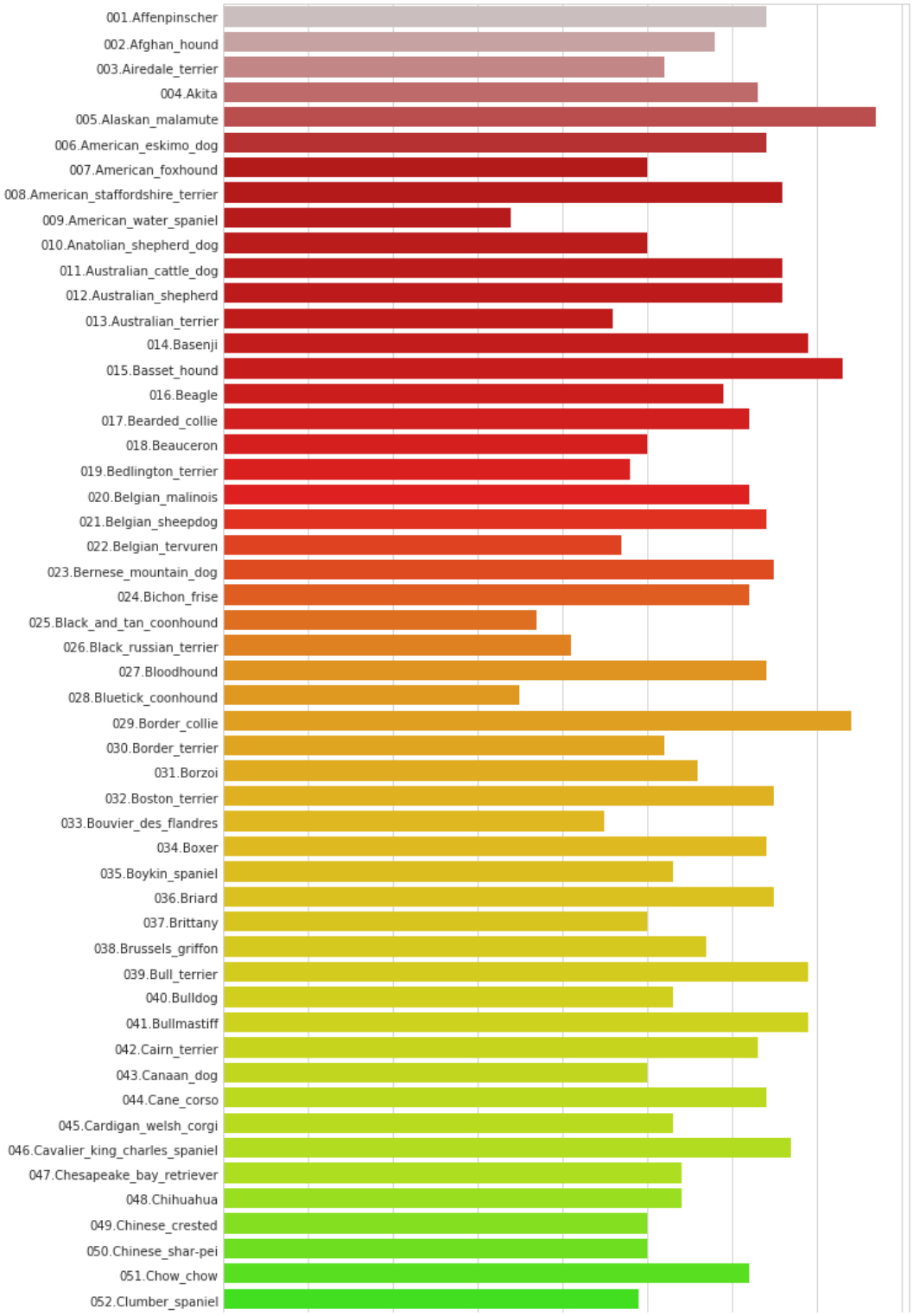
Datasets and Inputs

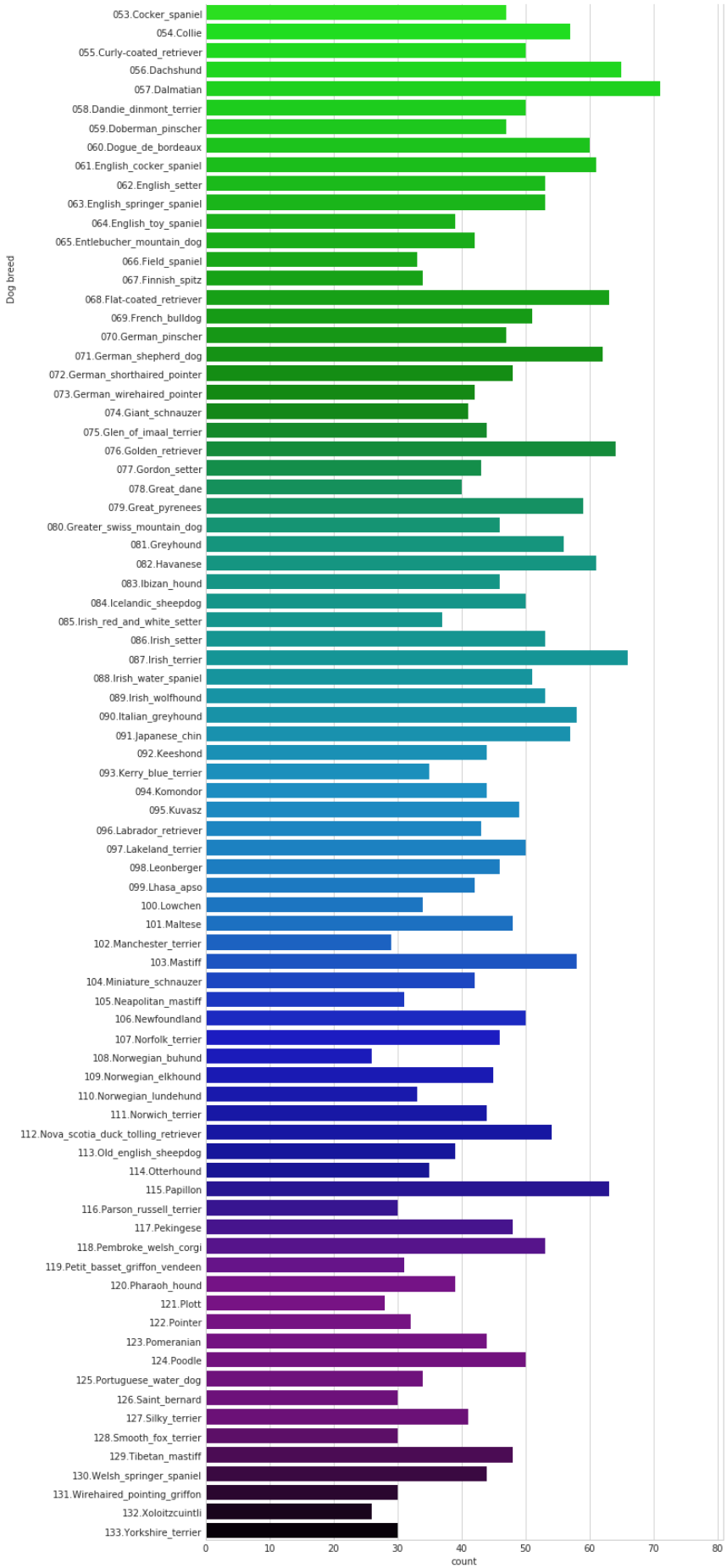
The dataset has image of 133 breeds of dogs. It a training set and a test set of images of dogs. Each image has a filename that is its unique id.

Source: https://www.kaggle.com/c/dog-breed-identification/overview/description

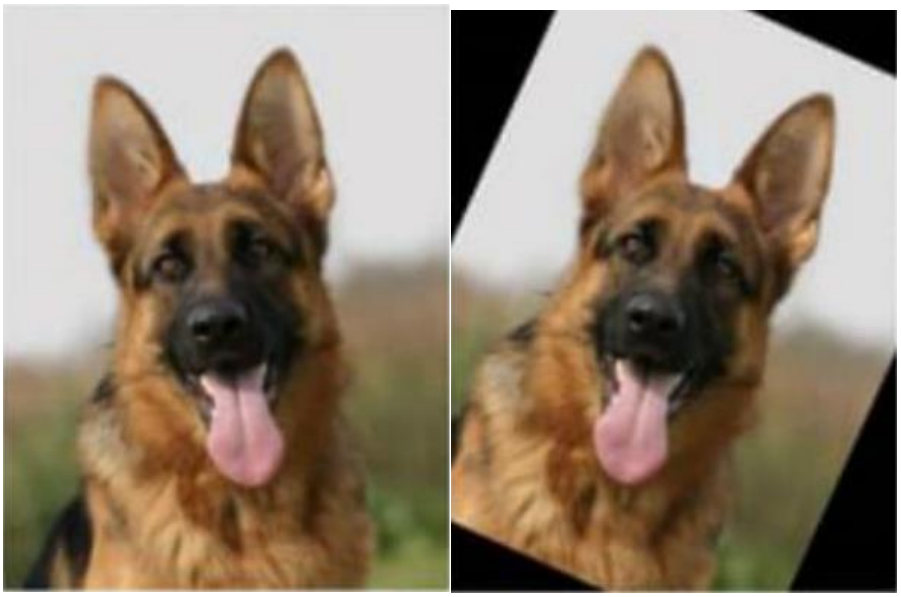
Here are some details about the dataset:

- Number of dog images in the dataset
  Total Files in /data/dog_images/train/: 6680
  Total Files in /data/dog_images/test/: 836
  Total Files in /data/dog_images/valid/: 835
- Number of classes in the dataset and the number of each classes is shown in the plot in the next 2 pages.

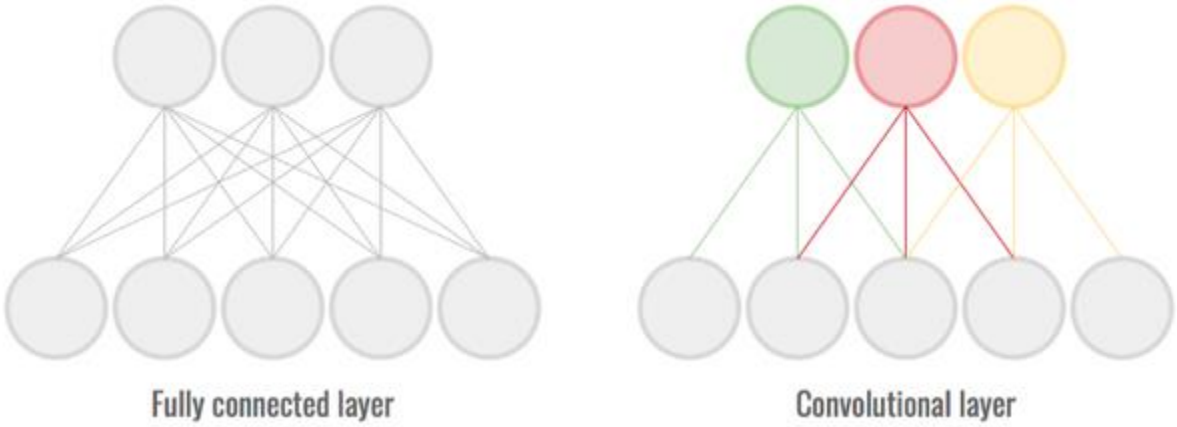

001.Affenpinscher
002.Afghan_hound
003.Airedale_terrier
004.Akita
005.Alaskan_malamute
006.American_eskimo_dog
007.American_foxhound
008.American_staffordshire_terrier
009.American_water_spaniel
010.Anatolian_shepherd_dog
011.Australian_cattle_dog
012.Australian_shepherd
013.Australian_terrier
014.Basenji
015.Basset_hound
016.Beagle
017.Bearded_collie
018.Beauceron
019.Bedlington_terrier
020.Belgian_malinois
021.Belgian_sheepdog
022.Belgian_tervuren
023.Bernese_mountain_dog
024.Bichon_frise
025.Black_and_tan_coonhound
026.Black_russian_terrier
027.Bloodhound
028.Bluetick_coonhound
029.Border_collie
030.Border_terrier
031.Borzoi
032.Boston_terrier
033.Bouvier_des_flandres
034.Boxer
035.Boykin_spaniel
036.Briard
037.Brittany
038.Brussels_griffon
039.Bull_terrier
040.Bulldog
041.Bullmastiff
042.Cairn_terrier
043.Canaan_dog
044.Cane_corso
045.Cardigan_welsh_corgi
046.Cavalier_king_charles_spaniel
047.Chesapeake_bay_retriever
048.Chihuahua
049.Chinese_crested
050.Chinese_shar-pei
051.Chow_chow
052.Clumber_spaniel


Dog breed
053.Cocker_spaniel
054.Collie
055.Curly-coated_retriever
056.Dachshund
057.Dalmatian
058.Dandie_dinmont_terrier
059.Doberman_pinscher
060.Dogue_de_bordeaux
061.English_cocker_spaniel
062.English_setter
063.English_springer_spaniel
064.English_toy_spaniel
065.Entlebucher_mountain_dog
066.Field_spaniel
067.Finnish_spitz
068.Flat-coated_retriever
069.French_bulldog
070.German_pinscher
071.German_shepherd_dog
072.German_shorthaired_pointer
073.German_wirehaired_pointer
074.Giant_schnauzer
075.Glen_of_imaal_terrier
076.Golden_retriever
077.Gordon_setter
078.Great_dane
079.Great_pyrenees
080.Greater_swiss_mountain_dog
081.Greyhound
082.Havanese
083.Ibizan_hound
084.Icelandic_sheepdog
085.Irish_red_and_white_setter
086.Irish_setter
087.Irish_terrier
088.Irish_water_spaniel
089.Irish_wolfhound
090.Italian_greyhound
091.Japanese_chin
092.Keeshond
093.Kerry_blue_terrier
094.Komondor
095.Kuvasz
096.Labrador_retriever
097.Lakeland_terrier
098.Leonberger
099.Lhasa_apso
100.Lowchen
101.Maltese
102.Manchester_terrier
103.Mastiff
104.Miniature_schnauzer
105.Neapolitan_mastiff
106.Newfoundland
107.Norfolk_terrier
108.Norwegian_buhund
109.Norwegian_elkhound
110.Norwegian_lundehund
111.Norwich_terrier
112.Nova_scotia_duck_tolling_retriever
113.Old_english_sheepdog
114.Otterhound
115.Papillon
116.Parson_russell_terrier
117.Pekingese
118.Pembroke_welsh_corgi
119.Petit_basset_griffon_vendeen
120.Pharaoh_hound
121.Plott
122.Pointer
123.Pomeranian
124.Poodle
125.Portuguese_water_dog
126.Saint_bernard
127.Silky_terrier
128.Smooth_fox_terrier
129.Tibetan_mastiff
130.Welsh_springer_spaniel
131.Wirehaired_pointing_griffon
132.Xoloitzcuintli
133.Yorkshire_terrier
0
10
20
30
40
50
60
70
80
count

- We will be using the following data augmentation on the dataset:
  - Resize to 224
  - Center Crop to (224,224)
  - Random Horizontal Flip
  - 
  - Random Rotation of 10 degrees

- 
- Normalization with mean=[0.485, 0.456, 0.406] and std=[0.229, 0.224, 0.225]

Solution Statement



Source: https://www.quora.com/What-are-the-advantages-of-a-convolutional-neural-network-CNN-compared-to-a-simple-neural-network-from-the-theoretical-and-practical-perspective

- The solution will be implemented using the convolutional neural networks. Convolutional neural networks help in capturing the special features of a data hence it is the best way to implement any image classification. CNN is also computationally efficient compared to fully connected layers.
- The solution will also use the pooling layers which further chooses important features only and reduces the training time.
- Relu will be used as the activation for hidden layers.
- Batch Normalization is used to improve the model convergence.
- Cross Entropy loss function is used for classification and it is the best estimator in terms of convergence rates.
- SGD optimizer of used since they are better for classification problems.

- Resnet50 will be used as a pretrained model. RESNET 50 allows a model to be deep without suffering the issues of vanishing gradient because of the skip connections. It also helps in communicating the features detected in top layers to the deeper layers which in turn helps in better learning of the model.

Benchmark Model

We are looking to understand how Convolutional Neural Networks can solve this problem. Hence the goals are tailored according to the resources and time available. Two model shall be implemented, one from scratch and the other using a pre-trained network. To prove that the model is learning features from the image we looking at the accuracy of at least 10% using the model created from scratch and 60% from the pre-trained network. The pre-trained model can be used to implement an interesting app.

Evaluation Metrics

For a classification problem accuracy will be a straight forward method to evaluation the model. Accuracy is the measure the number of correct decisions the classifier makes, divide by the total number of test examples. We could Recall, Precision, and F1- Score, but the accuracy will be enough for this problem statement.

Project Design

- The project will be implemented in Udacity workspace using an existing framework.
- GPUs will be used to train the model
- Pytorch will be used the implement the model
    - Dataloaders will be used to preprocess and fetch the data while training
    - One model will be created from scratch.
    - Another model will be implemented using a pre-trained network.
- A simple app will be created to use the model to predict the dog breed. It is made fun by using human images with the same model to match the human face to a dog breed.