CNN Project: Dog Breed Classifier

1.0 Problem Statement
The focus of this project is to write an algorithm for a dog classification application. The algorithm will be deployed within the application to solve the problem of dog identification i.e. given a user-supplied image, if a dog is detected in the image, it provides an estimate of the dog's breed or if a human is detected, it provides an estimate of the dog breed that is most resembling. This algorithm will potentially be based on a convolution neural network classifier that will perform the classification.

2.0 Domain Background
The concept of computer vision will be used in attempting to solve the problem of classifying the dog and human images. In computer vision, the objective is to describe the world in one or more images and to reconstruct its properties, such as shape, illumination, and color distributions. [1]
In particular, we shall explore the concept of convolutional neural networks, that use a hierarchical structure to 'learn' image features, starting with low-level features in the first hidden layer, then assemble them into larger higher-level features in the next hidden layer and so on across the depth of the layers. [2]

3.0 Datasets and Inputs
This project uses two datasets to build, train and test the algorithm that will be developed, namely:-

- Dog dataset, the dataset is obtained from Udacity AI Nano Degree course , it contains 8351 images of dogs. The dogs dataset will be used to train and test the model that will be built.
- Human dataset, this dataset is also obtained from the Udacity AI Nano Degree, it contains 13233 images of humans. The human dataset will be used to test to guide how the final algorithm is designed and also be used to test it.

4.0 Solution Statement
The solution is to build a convolutional neural network model trained on the data specified above to perform the dog breed classification within the dog application.

Given the small size of the dataset, transfer learning will be used so as to attain web/mobile application usable accuracy.Transfer learning involves the use of the lower hidden layers of an already existing pre-trained model that solves a similar task. This speeds up training considerably, and also works with limited data. [2]

5.0 Benchmark Model
The solution will be bench marked against the VGG16 state-of-the-art model. The VGG16 model achieves 92.7% top-5 test accuracy on ImageNet, a dataset containing over 14 million images belonging to 1000 classes. [3]

6.0 Evaluation Metrics
Given this is a classification problem, the evaluation will be based on the accuracy the model achieves on the test set, particularly, the model should achieve more than 60% accuracy on the test set.

7.0 Project Design
The solution is implemented in three steps:-

1. A human face detector is written to ascertain if a human face is contained in the given image, this returns a true or false response. To keep the solution manageable, an already existing face detection framework with proven performance will be used for instance the openCV face cascade.

2. A dog breed detector will be built in the project using the dog dataset, listed above in the section on datasets. To meet the accuracy specifications, transfer learning will be used in training the model, potential models to use include the state of the art VGG16. [3]

3. Finally, the algorithm will be designed as a combination of the two detectors listed above, the image is presented to a the face detector, and if a human face is detected, then the dog detector predicts the a dog breed of the most resemblance. If a human face is not detected, then the image is passed to just the dog detector and the dog breed detected. If the image is outside the listed scenarios, an error is returned.

References

[1] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag London Limited, 2011.

[2] Aurélien Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc., 2019.

[3] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.