**Capstone CNN Project: Dog Breed Classifier**

**Domain Background**

Image recognition is one of the exploding topics in machine learning [1-7]. By building and training state-of-the-art convolutional neural networks, machine can be used to classify different objects in the images. Inspired by ImageNet project [8], I will work on the convolutional neural networks project which can be used not only to recognize object in the image but also some interesting features about the recognized image.

**Problem Statement**

In this CNN project, I will build a pipeline to process the user-supplied images:

1. For an image of a dog, the algorithm will identify the predicted dog breed.
2. For an image of human, the algorithm will identify the resembling dog breed.
3. If neither is detected in the image, the algorithm will provide the output that indicates an error.

**Datasets and Inputs**

The datasets and inputs used in this project contains dog images and human images.

1. Dog images were provided by Udacity. There are 8351 dog images in total. There are 133 dog breed classes in the dataset, and the samples in each class ranges from 26 to 77. This is not a fully balanced data. Also, the image sizes are varied between images; therefore, resize is needed for the input of the CNN model.
2. Human images were downloaded from Labeled Faces in the Wild Project by the Computer Vision Laboratory established in the Computer Science Department at the University of Massachusetts. There are 13233 human images in total.

**Solution Statement**

In order to achieve the goal, I need to build three different machine learning models:

1. Human face detector
2. Dog detector
3. Dog breed predictor

If an image was detected as a human, dog breed predictor will be used to identify the resembling dog breed of this human.

If an image was detected as a dog, dog breed predictor will be used to predict its dog breed.

Otherwise, output that indicates an error will be provided.

**Benchmark Model**

To confirm the problem of dog breed identification is actually solvable, I will build a simple CNN benchmark model from scratch to classify the dog breed. The simple CNN model would be built by few layers of CNN followed by dense layers. The result will be the baseline score of the dataset I use.

**Evaluation Metrics**

There are 133 dog breed classes in the dataset, and the samples in each class ranges from 26 to 77. Since the data is not quite balanced, I will use not only use accuracy but also F1 score as evaluation metrics to see how good my dog breed classifier is.

**Project Design**

I will use Keras for this CNN project, which at least consists the following parts:

1. Build human detector
2. Build dog detector using pre-trained VGG16 model
3. Build dog breed classifier from scratch as benchmark model
4. Build dog breed classified using transferring learning
5. Write the algorithm to put everything together
6. Test my algorithm on several images

Some key pre-process steps for my CNN dog breed classifier:

1. Resize the images:

The image size of images varied. All images will be resized to 224x224.

1. Normalize the images:

Each image contains three channels, and the values in each channel ranges from 0 to 255. A normalization will be applied to each image by divided by 255.

1. Data augmentation:

There are only 8351 dog images, and there are 133 dog breed classes in the data. The sample numbers in each class ranges from 26 to 77. With the small number of dataset in each class, data augmentation will be used to improve the performance.

**Reference:**

1. Olga Russakovsky\*, Jia Deng\*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (\* = equal contribution) **ImageNet Large Scale Visual Recognition Challenge**. *International Journal of Computer Vision*, 2015.
2. J. Deng, O. Russakovsky, J. Krause, M. Bernstein, A. Berg, L. Fei-Fei. **Scalable multi-label annotation**. *ACM conference on human factors in computing (CHI)*, 2014.
3. O. Russakovsky, J. Deng, Z. Huang, A. Berg and L. Fei-Fei, **Detecting avocados to zucchinis: what have we done, and where are we going?**, *Proceedings of the International Conference of Computer Vision (ICCV). 2013*
4. J. Deng, A. Berg, K. Li and L. Fei-Fei, **What does classifying more than 10,000 image categories tell us?** *Proceedings of the 12th European Conference of Computer Vision (ECCV). 2010.*
5. O. Russakovsky and L. Fei-Fei, **Attribute Learning in Large-scale Datasets.** *Proceedings of the 12th European Conference of Computer Vision (ECCV), 1st International Workshop on Parts and Attributes. 2010.*
6. 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.*
7. J. Deng, K. Li, M. Do, H. Su, L. Fei-Fei, **Construction and Analysis of a Large Scale Image Ontology.** *In Vision Sciences Society (VSS), 2009.*
8. ImageNet project, <http://image-net.org/index>