

Capstone Proposal

Dog breed classification by Sai Aravind Sreeramadas

Domain Background

It was in 2012 , that alexnet which is a five layer cnn achieved a error rate very less than the stagnating error rate of image processing algorithms on the imagenet dataset.This showed that cnns were doing very well in terms of classification.Then came along many cnn models to beat that error rate every year with many modifications to their architecture and implementation.

Convolutional neural networks are the current leading top scorers in the image classification task. So, it is easy to differentiate the different categories i.e, dissimilar classes than the similar looking classes. For example dog and cat is easily differentiable but the types of aeroplanes or the different kinds of dogs.

This is tough because there exists similarities between the classes.Research has been done on this type of thing like using part localisation, part attention , landmarks e.t.c

Problem Statement

Given the dog breed data set of 10,222 images with 120 classes I have to build a cnn network which has a good validation accuracy and minimal log loss on the given dataset. After getting a good accuracy, a trade off between the model size and accuracy has to be made to deploy it on the Android platform.

Datasets and Inputs

The dataset to be used is the dog breed dataset which is hosted by the kaggle.This dataset is a subset of Stanford dog dataset.This consists of 10222 images of dogs with 120 classes. The class distribution of images is not uniform data.

This data is relevant as it has dogs which are similar in some aspects .ie, it is not a simple binary classification of two completely different classes , instead it is a fine grained classification.

In this data we use 9,200 samples for training and 1022 samples for validation.The training and validation set have the same class distribution.

Solution Statement

As I have less data instead of going for the scratch model, I am going to use the transfer learning approach .i.e, extract features from the xception , inceptionv3 and inception resnet v2 and concatenate these features and then build a dnn on this.

Benchmark Model

The vgg model has the performance of 78% on the test set of the Stanford dog dataset. Inceptionv3 has a accuracy of 93.

Evaluation Metrics

The evaluation metrics that we are going to use are the validation accuracy which tells us how our classifier is performing.

The other metric would be the logloss .this is used to tell with what confidence are we saying that our answer is right.

Project Design

- 1)download the kaggle dog dataset
- 2) now load the dataset using the keras load image function
- 3) we can see that the mean size of the images is around 400
- 4) so let's load the images of size 400*400
- 5) now split the data in to train and validation
- 6) scale the data by dividing it by 255
- 7) see the performance of a scratch model cnn
- 8) now take a pre trained model and extract the bottle neck features and then build a dnn on it.
- 9) see which models give a good accuracy
- 10) take 3 pretrained models and concatenate the bottle neck features and then build a dnn on top it.
- 11) now to further improve the model, we add more data by adding the flipped and cropped images
- 12) we tune the hyper parameters
- 13) look at the confusion matrix to see which classes are being wrongly classified
- 14) to see if we can do something about the miss classification
- 15) as this is a big network for deploying on mobile
- 16)we see the performance of the mobilenet on this dataset
- 17) deploy the mobilenet on the Android platform