Machine Learning Engineer Nanodegree

Capstone Proposal

Benjamin Rouillé d'Orfeuil

December 26, 2017

Domain Background

Convolutional Neural Networks (CNNs) have successfully been applied in the field of image recognition. This class of models has proven to be incredibly efficient at classifying images and often outperforms other machine learning algorithms at this task. To illustrate, very accurate predictions can be achieved on the well known MNIST database of handwritten digits[1] and the CIFAR-10 dataset[2] using a simple CNN architecture. Note that the number of classes in these datasets is very low (10 digits for the MNIST database and 10 object categories for the CIFAR-10 dataset) and that each object is vey different from one another. It follows that very few convolutional layers are necessary to accurately classify the images.

Image recognition on real world images requires the development of deep CNNs. For instance, the models developed for the ImageNet Large Scale Visual Recognition Challenge[3] (ILSRVC) have tens of convolutional layers. The reason is that the classification task for this competition is very complex. Not only the number of classes is important but the difference between the object categories can also be very tenuous. For example, among the 1000 classes in the ILSRVC dataset, 118 categories are dog breeds. Thus, It is not surprising that the CNNs developed for this challenge needs a large number of convolutional layers with a large number of filters. To illustrate, the ResNet50 model developed by Microsoft that won the 2015 ILSRVC edition has 50 convolutional layers and is quite deep with about 165 layers.

Problem Statement

Datasets and Inputs

Solution Statement

Benchmark Model

Evaluation Metrics

Project Design

[1] The MNIST database is available at `http://yann.lecun.com/exdb/mnist/`. A quick analysis of this dataset can be found here.

[2] The CIFAR-10 dataset can be found at the following url: `https://www.cs.toronto.edu/~kriz/cifar.html`. Predictions on this dataset are presented here.

[3] The 2017 challenge is described on the ImageNet website: `http://image-net.org/challenges/LSVRC/2017/index` and on Kaggle: `https://www.kaggle.com/c/imagenet-object-localization-challenge`.