Capstone-Proposal

April 14, 2018

1 Machine Learning Engineer Nanodegree

1.1 Capstone Proposal

Sahil Chutani
April 5th, 2018

1.2 Using Deep-Learning for automatic classification of Furniture and Home Goods Images

1.2.1 Domain Background

In the previous project I worked on image classification with CIFAR-10 dataset as a part of Udacity Machine Learning Nano Degree Program. Interested in exploring image classification more I ventured on Kaggle to look for active competitions regarding image classification with emphasis on research.

On Kaggle I came across a research competition titled "iMaterialist Challenge (Furniture) at FGVC5" which is focused on fine grained visual categorization. The competition is perfect for me as it proceeds along the lines of what I learnt in the previous project. Additionally, I would be able to apply Advanced Deep Learning Techniques Convolutional Neural Network (CNN) to the problem.

1.2.2 Problem Statement

As shoppers move online, it'd be a dream come true to have products in photos classified automatically. But, automatic product recognition is challenging because for the same product, a picture can be taken in different lighting, angles, backgrounds, and levels of occlusion. Meanwhile different fine-grained categories may look very similar, for example, ball chair vs egg chair for furniture, or dutch oven vs french oven for cookware. Many of today's general-purpose recognition machines simply can't perceive such subtle differences between photos, yet these differences could be important for shopping decisions.

In this problem the challenge is to develop an algorithm that will help with an important step towards automatic product recognition to accurately assign category labels for furniture and home goods images.

1.2.3 Datasets and Inputs

I would borrow data from Kaggle (https://www.kaggle.com/c/imaterialist-challenge-furniture-2018/data). Data are in JSON format. There are three files namely, train.json, validation.json, and test.json. I will be using train.json and validation.json for building the Deep Learning Model. I will download images programmatically with the URLs provided.

The training dataset includes images from 128 furniture and home goods classes with one ground truth label for each image. It includes a total of 194,828 images for training and 6,400 images for validation. Train and validation sets have the same format as shown below:

```
{

"images" : [image],

"annotations" : [annotation],

}

image{

"image_id" : int,

"url": [string]

}

annotation{

"image_id" : int,

"label_id" : int

}
```

1.2.4 Solution Statement

As I stated in the title, I propose to use deep learning in order to find a solution for the problem at hand. More specifically I propose to use Boosted Convolutional Neural Network (http://vision.cornell.edu/se3/wp-content/uploads/2016/08/boosted-convolutional-neural-1.pdf) as a part of the solution. I shall use Tensorflow library for the implementation of the solution.

A Convolutional Neural Network(CNN) consists of Convolution Layer, Pooling Layer, and Fully Connected Layer. For pooling layer, I propose to use Max Pooling. Additionally I shall use Rectified Linear Unit (ReLU) as activation funcion.

Taking reference from the above link, it has been shown that Boosted Convolutional Neural Network(BoostCNN) out performed other state-of-the-art methods when the performance comparison was made on bird classification using CUB200 dataset (http://www.vision.caltech.edu/visipedia/CUB-200-2011.html), aircraft classification using

FGVC-aircraft dataset (http://classif.ai/dataset/fgvc-aircraft/) and out performed all except Krause (http://vision.stanford.edu/pdf/joncvpr15.pdf).

For the project I propose to use BoostCNN as the algorithm.

Throughout the project I shall take referrence from the CIFAR image classification project I did previously in the Machine Learning Nano Degree program.

1.2.5 Benchmark Model

The problem I am trying to solve is a part of an open problem. It is a part of Fine-Grained Visual Categorization (FGVC5). Fine Categorization is one of the most interesting and useful open problems. For the current project I would use a Convolutional Neural Network (CNN) as the benchmark model.

1.2.6 Evaluation Metrics

The Boosted Convolutional network would trained in the same manner as the Convolutional Neural Network, therefore for Evaluation Metric I will be using Classification Accuracy which is given as :

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions\ Made}$$

1.2.7 Project Design

The first step in this project would be data collection. For Step 1 in collection of data I shall download the training files and validation files from Kaggle Page (https://www.kaggle.com/c/imaterialist-challenge-furniture-2018#description). In the next step for collection of data I shall download the images programmatically and save them in a file.

As the next step in the project I shall perform univariate, bivariate, and multivariate visualizations. Additionally I will perform data cleaning if necessary.

Given the success of CNNs I shall implement the benchmark Convolutional Neural Network Model (CNN). I will use the benchmark model to compare the performance of the proposed Boosted Convolutional Neural Network Model (boostCNN). I will utilizing layers of convolutions, max pooling for pooling layer, Rectifier Linear Units (ReLU) as activation function and finally applying fully connected layers to achieve the benchmark model.

Fourth, I shal implement the proposed boostCNN model, and compare its performance against the benchmark model. Comibining Boosting with Decision Trees has shown better performance in comparison with Decision Trees alone. Moving along the same line for inspiration to adding boosting to CNNs. In this project I will endeavor to combine the benefits of boosting with CNNs and try to achieve a model that out performs the one with CNN only.

Finally, I shall write a succint report comparing the performance of the two models.

1.2.8 References

1. https://www.kaggle.com/c/imaterialist-challenge-furniture-2018#description
2. https://vision.cornell.edu/se3/boosted-convolutional-neural-networks/
3. https://vision.cornell.edu/se3/wp-content/uploads/2016/08/boosted-convolutional-neural-1.pdf
4. http://vision.stanford.edu/pdf/joncvpr15.pdf
5. https://cilvr.cs.nyu.edu/diglib/lsml/lecture03-trees-boosting.pdf
6. https://arxiv.org/pdf/cs/0502006.pdf