

Machine Learning Engineer Nanodegree

Capstone Project

Benjamin Rouillé d'Orfeuil

May 17, 2018

I. Definition

Project Overview

[Yelp](#) is a social networking site that publishes crowd-sourced reviews about local businesses. About two years ago, Yelp challenged Machine Learning practitioners to build a model that automatically tags restaurants with multiple labels using a dataset of user-submitted photographs. The goal of this project is to develop such a model.

The competition was hosted by [Kaggle](#), a platform where data scientists use their skills to produce the best models for predicting and describing datasets uploaded by companies and users. The various datasets and inputs are accessible via the Yelp Restaurant Photo Classification competition webpage¹. Yelp provides a training dataset (234,842 photographs) and a test dataset (237,152 photographs) for this competition². Each photograph belongs to a business and the task is to predict the business attributes purely from the business photographs. There are 9 different attributes in this problem:

- good for lunch;
- good for dinner;
- takes reservations;
- outdoor seating;
- restaurant is expensive;
- has alcohol;
- has table service;
- ambience is classy;
- good for kids;

and a total of 2,000 (10,000) businesses in the training (test) dataset. Note that this is a multi-instance multi-label classification problem. Each business has multiple photographs and predictions need to be made at the business level. Also, multiple labels can be assigned to each business meaning the classifier needs to account for label dependencies.

Problem Statement

Neural networks have proven to be incredibly efficient at classifying images and often outperform other machine learning algorithms at this task. It comes then as no surprise that deep learning models are used extensively in this project. One now faces two options: i) build and train a deep neural network

¹See <https://www.kaggle.com/c/yelp-restaurant-photo-classification>

²The datasets are quite large. Both the training and test archive files have a size of about 7 GB.

from scratch or ii) use transfer learning³. The properties of the dataset such as its size and nature usually dictate the type of approach to adopt. The Yelp dataset being both large and complex, it would be unrealistic to train a deep neural network model from scratch given this task would require fine expertise and enormous resources. Also, deep neural networks that have been pre-trained on large and diverse dataset like ImageNet⁴ captures universal features in its early layers that are relevant and useful for most computer vision problems. Thus, leveraging such features allows to reach a better accuracy than any method that would rely only on the available data. For all those reasons, transfer learning was a better approach for this project and a pre-trained state-of-the-art deep learning model has been used as a fixed feature extractor.

II. Analysis

III. Methodology

IV. Results

V. Conclusion

³Machine learning technique where a model trained on one task is re-purposed on a second related task.

⁴Large visual database designed for use in visual object recognition software research. Research teams have been evaluating their algorithms and competing for years using this large visual database.