Capstone Project Datasets

04.06.2019

Ky Pham

Springboard AIC, Cohort 02.04.2019

Riverside, CA, USA.

Overview

This document describes some image based datasets which will be used for the capstone project.

Capstone Project

As an immigrant, sometimes I had a hard time to find a talking point when hanging around with new friends, coworkers, and customers. I would like to build a mobile app leveraging the camera to take a picture of foods, drinks, movies, cars, signs, buildings, etc., then using a backend service to recognize it in order to provide interesting facts:

- What is the name
- Brief history: where it came from; when it was first made, etc.
- Who (famous people) have tried (foods/drinks), stared (movies), owned (cars), lived (buildings), etc.

This information can be shown as text on the screen or streamed as speech to earplug so the person can have some ideas to engage a conversation.

The app can be extended to salespersons to personalize welcoming.

Here are some datasets to begin with:

- Foods (https://www.vision.ee.ethz.ch/datasets_extra/food-101/)
  Download: https://s3.amazonaws.com/fast-ai-imageclas/food-101.tgz
- Cars (https://ai.stanford.edu/~jkrause/cars/car_dataset.html):
  Download: https://s3.amazonaws.com/fast-ai-imageclas/stanford-cars.tgz
- Movies (REST API):
  http://www.omdbapi.com

Some ML techniques:

- Classification/Deep Learning: YOLO, Caffe
- Text Summarization (Subject extraction using Google API)
- Text-to-speech generation

How to collect facts:

- Google search "<term> fun facts" and obtain the first match (pressing next button for another down on the list)

Dataset 1 - Foods 101

(https://www.vision.ee.ethz.ch/datasets_extra/food-101)

This dataset is more than 5GB. It consists of 101 food categories with 101,000 images. For each category, there are 250 manually reviewed test images and 750 training images. On purpose, the training images were not cleaned, and thus still contain some amount of noise. This comes mostly in the form of intense colors and sometimes wrong labels. All images were rescaled to have a maximum side length of 512 pixels.

Dataset 2 - TBD