DATS-6103: Data Mining Final Project Group Proposal

Popular Attraction/Landmark Recognition Using Google Landmark Dataset

With a rapid increase in the use of smartphones and other social apps, Image Recognition, Image Classification and Image Processing are the latest concepts that interest data engineers in computer vision tasks. A major challenge with image classification is the lack of a large, annotated dataset to train better and robust models.

Problem Statement:

Recognizing and training the model to identify any landmark is a challenging task as the appearance of the landmark varies with geometry, illumination and a different aspect ratio of the image presented. To overcome this issue, a collection of images is used to capture typical appearance of the location. This project will focus to build a model that recognizes a given popular attraction or landmark using Google landmark dataset. This landmark recognition model will be handy to identify the name of a landmark in the image. This will also helpful for photo organization in smartphones and fields like aviation, maps, crime - solving, etc.

Dataset:

In order to capture the typical appearance of an image via a collection of images, we need a large annotated landmark dataset. Google has released its latest landmark dataset named, Google-Landmarks-v2 (September 2019) which makes it our ideal choice for landmark recognition and retrieval purposes. This dataset includes over 5 million images with more than 200,000 diverse landmark classes. Google has published this dataset in 3 sets – train, index and test. The train and test files are used for landmark recognition and index file is used for retrieval purposes. Train dataset consists of image details of various landmarks, while test dataset consists of images that include no landmark, one landmark or multiple landmark. The major challenge while using this dataset is that of a highly imbalanced training dataset. This is because since there are large number of categories, also many classes with single digit training data which makes it difficult to classify and train the model for such classes.

Train dataset – 4132,914 location data with 203,094 unique classes

Test dataset – 117,577 data points

Since the dataset is highly imbalanced, performing data pre-processing needs to be considered before training the model. The dataset also needs to be cleaned to find any broken url (analyzing the image). The dataset is created by crowdsourcing the landmark available online. Each image might have different pixel size; hence these images need to be resized to one uniform pixel size for analysis and training.

Data Mining Algorithm & Network:

A HOG (Histogram Oriented Gradient) is a feature descriptor that is used to extract features from the images and thereby feature vectors are created. Further, we aim to use SVM (Support Vector Machine) classifier for this project. The dataset is labeled, however high dimensional feature spaces are needed to handle it, hence, SVM generally gives the best performance in such cases. With the help of customized machine learning packages, we plan to code and train the dataset in Python programming language, hence latest Python version and any IDE will be used to run code.

Reference Materials:

1. Announcing Google-Landmarks-v2: An Improved Dataset for Landmark Recognition & Retrieval (2019, September),
   Retrieved from: https://ai.googleblog.com/2019/05/announcing-google-landmarks-v2-improved.html
2. The Common Visual Data Foundation(2019, September), Google Landmarks Dataset v2,
   Retrieved from: https://www.kaggle.com/c/landmark-recognition-2019
3. Y. Li, D. J. Crandal and D. P. Huttenlocher, Landmark Classification in Large-scale Image Collections,
   Retrieved from: https://www.cs.cornell.edu/~yuli/papers/landmark.pdf
4. A. Crudge, W. Thomas and K. Zhu, Landmark Recognition Using Machine Learning,
   Retrieved from:
   http://cs229.stanford.edu/proj2014/Andrew%20Crudge,%20Will%20Thomas,%20Kaiyuan%20Zhu,%20Landmark%20Recognition%20Using%20Machine%20Learning.pdf
5. Y. Takeuchi, P. Gros, M. Hebert and K. Ikeuchi, Visual Learning for Landmark Recognition,
   Retrieved from: https://www.cs.cmu.edu/~takeuchi/iuw97/iuw97.html

Performance Metrics:

We aim to use the most common performance metrics utilized in image classification such as accuracy score, confusion matrix, precision, sensitivity (Recall), specificity, AUC curve, F1 score, etc.

Rough Schedule:

We intend to complete this project in a span of 5 weeks (by Dec 1st, 2019). Below is an estimate of our schedule over the next few weeks.

- ❏ Dataset Finalization – 1 week
- ❏ Understanding the dataset, Proposal Draft, Git set up, Action plan – 3 Days
- ❏ Data Preprocessing – 1 Week
- ❏ Data Training & Prediction using SVM – 1 Week
- ❏ Performance analysis – 5 Days
- ❏ Documentation (Group & individual Report, PPT) – 1 Week