Google Landmark Recognition

Kaggle Competition

CMPE-255
Term Project Report
San Jose State University
By
Anusha Velumani, Sahana, Purva Deekshit, Sachin Guruswamy, Nikhil Saunshi

Abstract:
The project aims to identify various landmarks across the globe using the input image. The supervised deep learning model is able to predict landmark labels using the given input image.

Dataset:
Dataset consists of 5,000,000 rows containing URLs for images along with their landmark IDs as labels.

https://www.kaggle.com/c/landmark-recognition-2019/data

The dataset consists of approximately 200K classes of images, categorized according to type of images. However, number of images per class is very less for many classes.

Pre-processing:
Images are normalized and resized to 299 x 299 resolution. We have performed image augmented by flipping the images to left, right, upwards and downwards, in order to increase the number of images per class.

Approach:
We have bootstrapped the model with pre-trained weights from Google's Xception model, which is an interpretation of GoogLeNet Inception V3 module. It performs depth wise convolution followed by 1 x 1 pointwise convolution at each hidden layer, to extract details of each pixel in 3D. However, the training is done in two separable mappings of 2D + 1D. This makes the process of learning more efficient and the model learns only informative features from the image. For activation function, we have used a generalized mean pooling, so that the algorithm will learn only informative features.

Train and Test:
Out of 5,000,000 images, we are able to download 200,000 images, which we split in training and validation dataset.

Result:
After running the algorithm for 50 epochs with a batch size of 48, we were able to achieve 22% accuracy because of large dataset and not enough images to learn features from all classes.

Future Scope:
Currently, with existing system specifications and algorithms, the highest achieved accuracy is 60%. We are trying to download complete dataset of ~5 million images using high processing GPUs to train the algorithm, in order to achieve more accuracy.

References:
[1] http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8451317&isnumber=8451009

[2] https://medium.com/@abhinaya08/google-landmark-recognition-274aab3c71ae

[3] https://arxiv.org/pdf/1711.02512.pdf

[4] https://cs230.stanford.edu/projects_spring_2018/reports/8291223.pdf

[5] http://help.clarifai.com/api/batch-processing/batch-processing-with-python

[6] https://towardsdatascience.com/google-landmark-recognition-using-transfer-learning-dde35cc760e1

[7] https://www.kaggle.com/c/landmark-recognition-challenge/discussion/58050#latest-340538