Kaggle Competition

--iMet Collection 2019

04.10.2019

Competition URL: https://www.kaggle.com/c/imet-2019-fgvc6

Xiangyu Chen
Lei Liu
Yuchen He

Final Project : INFO 7374 Special Topics in Info Systems

Overview

iMet Collection - FGVCx competition based on The Met's digitized collection

The Metropolitan Museum of Art in New York, also known as The Met, has a diverse collection of over 1.5M objects of which over 200K have been digitized with imagery. The online cataloguing information is generated by Subject Matter Experts (SME) and includes a wide range of data. These include, but are not limited to: multiple object classifications, artist, title, period, date, medium, culture, size, provenance, geographic location, and other related museum objects within The Met's collection. While the SME-generated annotations describe the object from an art history perspective, they can also be indirect in describing finer-grained attributes from the museum-goer's understanding. Adding fine-grained attributes to aid in the visual understanding of the museum objects will enable the ability to search for visually related objects.

Images

Multiple modalities can be expected and the camera sources are unknown. The photographs are often centered for objects, and in the case where the museum artifact is an entire room, the images are scenic in nature.

Annotations

Each object is seen by a single worker without a verification step. Workers are advised to add multiple labels from an ontology provided by The Met, and additionally are allowed to add free-form text when they see fit. The crowd is able to view the museum's online collection pages and is advised to avoid annotating labels already present. Specifically, the crowd is advised to annotate labels related to what they "see" or what they infer as the object's "utility." We consider these annotations noisy.

Goals

1. To explore the distribution of attributes for the Metropolitan Museum's collection.

2. To make a satisfied predictive classification for each arts and hopefully get the team become the top 30% among all the competitors.

3. Further using the model for implementing some other potentially relative classification software, such as photo tagging or recommending system.

Data

iMet Collection 2019-FGVC6
https://www.kaggle.com/c/imet-2019-fgvc6/data

Each data sample contains one image and at least one attribute label from a label set of 1103 attributes.

The filename of each image is its *id.*

Labels

1103 include cultures & tags

labels.csv provides descriptions of the attributes

Training set

109,274 Samples

train.csv gives the *attribute_ids* for the train images in /train

Validation set

7,443 Samples

Test set

38,814 Samples

test contains the test images. You must predict the *attribute_ids* for these images.

Metric

F_beta-score on the test set will be used as final score.

Evaluation

F2 score:

Recall is the ratio of true positives to all actual positives (tp + fn). The F2 score is given by:

$$\frac{(1+\beta^2)pr}{\beta^2 p + r} \quad \text{where} \quad p = \frac{tp}{tp+fp}, \quad r = \frac{tp}{tp+fn}, \ \beta = 2.$$

The mean F2 score is formed by averaging the individual F2 scores for each id in the test set.

Process Outline

1. Data Preprocessing
   - Data Cleaning, handling missing values
   - Resize image size
   - Handle 1103 types of labels
2. Exploratory Data Analysis
3. Select Best Model and performance

Milestones

Timeframe	Delivery
Day 1-4	Data Preprocessing and Exploratory Data Analysis
Day 5-10	Model Building, Training
Day 11-12	Model selection and Performance
Day 13-14	Documentation

Deployment Details

1) Language: Python

2) Available pre-trained model we might use:

VGG19 model [5]

ResNet model [6]

DenseNet model [7]

Neural Architecture Search Network (NASNet) models [8]

3) Tools for Analysis: Jupyter Notebook, Google Cloud Platform

Reference and Sources:

[1] Vladimir Iglovikov, Alexey Shvets. TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation. 2018
https://arxiv.org/abs/1801.05746

[2] Vladimir Iglovikov, Sergey Mushinskiy, Vladimir Osin. Satellite Imagery Feature Detection using Deep Convolutional Neural Network: A Kaggle Competition. 2017
https://arxiv.org/abs/1706.06169

[3] Haiyong Zheng, Ruchen Wang, Zhibin Yu, Nan Wang, Zhaorui Gu and Bing Zheng. Automatic plankton image classification combining multiple view features via multiple kernel learning. 2017
https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-017-1954-8

[4] Mario Lasseck. Bird Song Classification in Field Recordings: * Winning Solution for NIPS4B 2013 Competition. 2013
http://www.animalsoundarchive.org/RefSys/Nips4b2013NotesAndSourceCode/WorkingNotes_Mario.pdf

[5] Deep Residual Learning for Image Recognition
https://arxiv.org/pdf/1512.03385.pdf

[6] Very Deep Convolutional Networks for Large-Scale Image Recognition
https://arxiv.org/pdf/1409.1556.pdf

[7] Densely Connected Convolutional Networks
https://arxiv.org/pdf/1608.06993.pdf

[8] Learning Transferable Architectures for Scalable Image Recognition
https://arxiv.org/pdf/1707.07012.pdf