Landmark Detection

Jayanth Dasari
2019BCS-016
ABV-IIITM, Gwalior

Satya Pavan Kalyan Vemula
2019BCS-069
ABV-IIITM Gwalior

Mahesh Bommisetty
2019BCS-013
ABV-IIITM Gwalior

Anunay Nalam
2019BCS-034
ABV-IIITM Gwalior

I. Problem Statement

The Landmark Detection project aims for Instance Level Recognition, to detect the Landmarks inside given images with different perspectives, from various camera angles and lighting conditions. This technology can predict landmark labels directly from image pixels to help people better understand and organize their photo collections.

II. Current Research Area

In this domain, the most popular dataset, the Google Landmarks Datasetv2, consists of over 5M images and over 200k distinct instance labels, making it the largest instance recognition dataset to date[4]. Apart from GLDv2, the Oxford and the Paris datasets[3] are other famous datasets. In image detection/recognition problems, there are mainly three methods, Basic Recognition, Fine-Grained Recognition involving distinction of species/models/styles, and Instance Level Recognition. Currently, the main focus is on Instance Level Recognition and Google Landmark Recognition Dataset due to the vast number of included images and classes.

III. Motivation and Justification

Landmark Recognition has many applications like Visual Search for recognizing famous buildings, Personal Photo Search for organizing one's image library, grouping photos taken at a landmark, and gaming through augmented reality. This project has vast datasets, including many images that are distractors, not landmarks, and have much larger number of classes. Solving these problems is a huge challenge that motivated us to take up this project as our course project.

IV. Literature Survey

Cheng Xu et al.[5] developed their model on the Google Landmark detection dataset using features and classification logits further optimized with an ArcFace Loss. Included an efficient pipeline for re-ranking predictions by adjusting retrieval scores and got a 0.489 score with the ensembled model.

Nilwong et al.[2] presented a method alterantive to conventional ways for outdoor localization relying on Faster RCNN and the feedforward neural network (FFNN) trained on their custom dataset images with geotags and labeled bounding boxes of the koganei campus of Hosei University.

Christof Henkel et al.[1] developed an end-to-end instance level recognition method for labeling and ranking landmark images. In their method, they embed images in a high-dimensional feature space, classify them based on visual similarity, re-ranking predictions, and filter noise based on their similarity to out-of-domain images.

V. Work Done

- IIITM Campus dataset collection, Collected the IIITM Campus images under different perspectives, lighting conditions, and from various angles.
- Collected the campus images data by considering 14 prominent landmarks. Custom Dataset contains 500 images with intra-class variability, equally distributed.
- Explored various SOTA models(Resnet, Faster RCNN, VGG, MobileNet and soon) and techniques on Google Landmark Detection dataset and other similar datasets.

VI. Future Work

- Finetuning the pre-trained SOTA models to the custom dataset.
- Exploring the other pre-trained models, augmentation methods, and data processing methods.
- Expanding Dataset.

VII. Expected Deliverables

A Final report explaining our approach and the analysis we did before finalizing the approach, the results of our approach, and the trained models.

References

[1] Christof Henkel and Philipp Singer. *Supporting large-scale image recognition with out-of-domain samples*. 2020. arXiv: 2010.01650 [cs.CV].

[2] Sivapong Nilwong et al. "Outdoor Landmark Detection for Real-World Localization using Faster R-CNN". In: Oct. 2018, pp. 165–169. DOI: 10.1145/3284516.3284532.

[3] Filip Radenović et al. *Revisiting Oxford and Paris: Large-Scale Image Retrieval Benchmarking*. 2018. arXiv: 1803.11285 [cs.CV].

[4] Tobias Weyand et al. "Google Landmarks Dataset v2 - A Large-Scale Benchmark for Instance-Level Recognition and Retrieval". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.

[5] Cheng Xu et al. *3rd Place Solution to Google Landmark Recognition Competition 2021*. Oct. 2021.