Image Annotation Based Information Retrieval
Jnanesh D, Sachin and Prof. Surabhi Narayan
Department of Computer Science and Engineering, PES University

Problem Statement

Image annotation-based information retrieval is a full stack application which detects the objects in an image through bounding boxes and provide appropriate information about that object through Google Vision API.

Background

Object detection is widely used in the field of computer vision and crucial for variety of applications, e.g., self-driving car. Object detection methods have been continuously developed, and generated numerous approaches which obtained promising achievements. Several machine learning models have been developed for the same. We incorporated SSD for object detection model which has state of the art performance.
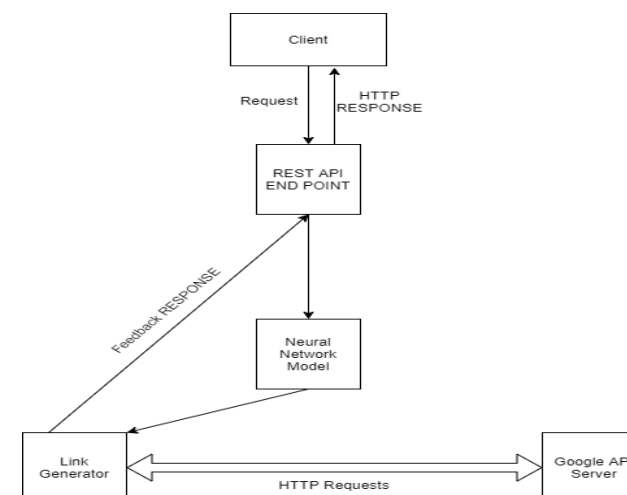
Dataset and features

In this project we have used MS COCO Dataset which has over 3,20,000 images which labelled with 80 different classes of images. SSD is trained on MS COCO 2014 dataset.

Features: In this project we have a client server architecture where a user can either upload an image or capture one through the integrated camera in a device and upload it to the user. Server receives image and responds with the objects detected. Upon clicking an object additional request is sent to receive the information about object in particular.

Design Approach / Methods

We have a client server architecture – 2 servers and a client. Client-side UI will request the SSD server by uploading the image and server responds the objects detected. Upon clicking on each object, a request is sent to information server and receives information about that image along with the appropriate links.

-: System Design :-

Results and Discussion

Upon evaluating our model on validation set over 60,000 images, we have obtained

| Metric	IoU	Value
Avg Precision	0.5:0.95	0.47
Avg Precision	0.5	0.556
Avg Recall	0.5:0.95	0.3
Avg Precision	0.7	0.378
Avg Precision Sm	0.5:0.95	0.116
Avg Precision M	0.5:0.95	0.404

Summary of Project Outcome

SSD and YOLO models are able to detect maximum possible objects from image with the help of google vision API we are able to retrieve useful links and with the help of wiki search we provide description about the objects.

System input: System Output
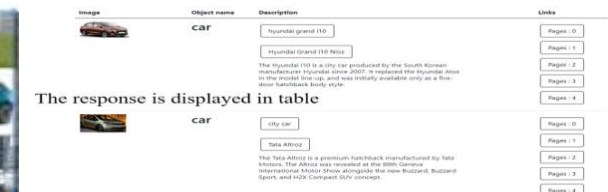
Conclusion and Future work

Object recognition Image Annotation Based Information is a challenging and exciting task of computer vision pushing the limits of computational capabilities. Our project is one such attempt to build a full stack application which not only detect the objects in an image but also provide the information about the object present in the image.
Future work:
1. Deploying it into cloud to make it available to use.
2. Implementing non-functional requirements like load balancing, scalability and security.

References

[1] Joseph Redmon, Santosh Divvalay, Ross Girshick, Ali Farshadi "You Only Look Once: Unified, Real-Time Object Detection" 2016

[2] Wei Liu, Dragomir Angelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng Yang Fu, Alexander C Berg "SSD :Single Shot MultiBox Detector" 2016

[3] A Practical Study About the Google Vision API - Daniel Pedro Ferreira Lopes,Antonio J. R

PW21SBN05 - Team

Prof. Surabhi Narayan	Jnanesh D	Sachin