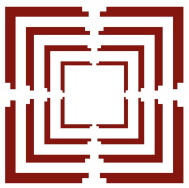

Ahmedabad
University

Machine Learning + Computer Vision Project

Group - 11

Week-4: Progress Report

Project title:

Evaluate performance of various object detection techniques (in case of small objects) on AU Drone dataset.

Group Members:
AU2040014 Jay Patel
AU2040021 Dhanya Mehta
AU2040265 Spandan Shah

Task performed in this week:

We have started to understand the sysNet model and End-to-End object detection model with transformer using DEtection TRansformer.
The transformer model is a state-of-the-art approach to object detection in computer vision. Unlike traditional object detection methods, which rely on complex pipelines of handcrafted components such as feature extractors, region proposal networks, and classifiers, DETR directly predicts object bounding boxes and their corresponding class labels from a full image, without relying on any predefined object proposals.

DETR consists of a backbone convolutional neural network (CNN) that extracts features from the input image and a transformer network that performs object detection. The transformer network uses a set of learned queries to attend to different regions of the image features and generates a set of object queries. These queries are matched with the features to predict the class and location of each object in the image. The End-to-End nature of DETR makes it simple and efficient, allowing for fast training and inference times. Additionally, it achieves state-of-the-art results on standard benchmarks for object detection, surpassing traditional methods like Faster R-CNN and Mask R-CNN.

The main goal of SyNet is to address the problem of large-scale graph representation learning by bridging and synthesizing communities within the graph. SyNet achieves this by using a two-stage approach: first, it identifies communities within the graph and then synthesizes them into a more compact representation. In the first stage, SyNet uses a clustering algorithm to identify the communities. In the second stage, it uses a graph neural network (GNN) to synthesize the communities into a compact representation.

The authors of the paper evaluated SyNet on several benchmark datasets and showed that it outperforms state-of-the-art methods for large-scale graph representation learning. They also demonstrated that SyNet can be applied to various downstream tasks, such as link prediction and node classification, with competitive results. Overall, SyNet provides a promising approach to large-scale graph representation learning by leveraging the inherent structure of communities within the graph.

We have read and understood the paper thoroughly and are planning to implement it within the following week.

Problems faced during the week:
We have started the implementation of the SyNet algorithm and are in phase of solving errors before solving the training the model on the COCO dataset and after testing the model on the same, we will start to train the model on the VisDrone dataset as well.

The task to be performed in the next week:

- Solve all the errors for both the models and test it out for detection of small objects.
- Use vis-drone dataset, and compare the accuracy for the same for both the models.

References:

- https://arxiv.org/pdf/2012.12991.pdf
- https://github.com/mertalbaba/SyNet
- https://cocodataset.org/#download
- https://arxiv.org/pdf/2005.12872.pdf