Course Project Abstract

Project Title

A Comprehensive Evaluation of YOLO as a Single-Stage Object Detector in Comparison to R-CNN(Two-Stage Detector) and Vision Transformer(Transformer based Detector)

Group Members

Yoontae Park(yp2201), Yoobin Cheong(yc5206), Yeong Koh(yk2678)

Abstract

Object detection is a fundamental task in computer vision that involves identifying and localizing objects within an image or video. In recent years, several object detection algorithms have been developed, including single-stage detectors such as YOLO and two-stage detectors such as R-CNN, as well as recent transformer-based detectors like Vision Transformer.

This research project proposes a comprehensive evaluation of YOLO as a single-stage object detector in comparison to R-CNN and Vision Transformer. We will provide a detailed overview of each algorithm, including their architecture, training methodology, and inference procedure. We will then compare their accuracy, precision, recall, and F1-score on benchmark datasets, such as COCO, Pascal VOC, and KITTI, to assess their performance. As a potential extension, we might also consider the trade-off between speed and accuracy and compare the detection speed of each algorithm. This aspect is particularly important in real-time applications, where the detection speed can directly impact the performance of the system.

Overall, this project aims to provide a comprehensive analysis of state-of-the-art object detectors and help researchers and practitioners select the most appropriate algorithm for their specific use case.

Reference

1. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors: https://arxiv.org/pdf/2207.02696.pdf
2. YOLOv7: A Deep Dive into the Current State-of-the-Art for Object Detection: https://towardsdatascience.com/yolov7-a-deep-dive-into-the-current-state-of-the-art-for-object-detection-ce3ffedeeaeb
3. YOLOv7 Implementation Github: https://github.com/WongKinYiu/yolov7
4. Simple Open-Vocabulary Object Detection with Vision Transformers: https://arxiv.org/pdf/2205.06230.pdf
5. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks: https://arxiv.org/pdf/1506.01497.pdf