

Object Detection on Cityscapes

A. Project Description

a. Cityscapes Dataset

In this project, we will develop an object detection for the cityscape dataset. Cityscapes is a large-scale database that focuses on scene understanding of urban street scenes. It provides semantic, instance-wise, and dense pixel annotations for 30 classes grouped into 8 categories (flat surfaces, humans, vehicles, constructions, objects, nature, sky, and void). The dataset consists of around 5000 fine annotated images and 20000 coarse annotated ones. Data was captured in 50 cities during several months, daytime, and good weather conditions. It was originally recorded as video so the frames were manually selected to have the following features: a large number of dynamic objects, varying scene layouts, and varying backgrounds.

b. Object Detector

There are many deep learning-based object detectors, e.g., Faster R-CNN, YOLO, DETR, DINO, etc. In this project description, we will YOLO as an example of the detector. You are free to use any other object detector.

YOLO, or "You Only Look Once", is a popular object detection algorithm that uses a convolutional neural network (CNN) to predict bounding boxes and class probabilities of objects in input images. YOLO is fast and efficient, making it a good choice for real-time object detection. YOLO's approach is different from traditional methods, which involve multiple stages. Instead, YOLO accomplishes detection in a single forward pass through the neural network. In this project description, we will YOLO V7 to train an object detector on the Cityscapes dataset.

c. Training The Detector

First, you need to register an account to download the Cityscape dataset:

<https://www.cityscapes-dataset.com/>

Second, you will clone this repository: <https://github.com/truongthanhdatt/YOLO-V7-Cityscapes>

Third, you need to follow the instructions on the GitHub repository to test and train the YOLO V7 object detector.

B. Questions

- a. Train the object detector on the Cityscapes dataset

- b. Use your detector to detect objects on the videos provided in the Cityscapes dataset (leftImg8bit_demoVideo.zip). You can download the video from here:
<https://www.cityscapes-dataset.com/file-handling/?packageID=12>