Data Documentation for Object Detection and Segmentation Project

Introduction

This document provides a comprehensive overview of the dataset used in our object detection and segmentation project. The project utilizes the Cityscape dataset, along with additional images sourced from YouTube videos, to detect and segment objects in urban street environments. The original data can be accessed via the following drive link. This documentation covers the sources and preprocessing steps.

Dataset

The dataset comprises images from the Cityscape dataset and additional Frames from YouTube videos showing driving scenarios in German towns. The images have been annotated and preprocessed to ensure they are properly cleaned and formatted for the project.

Data Description

Sources:

Cityscape Dataset: This dataset contains urban street images with annotations for 18 different classes, primarily captured in German cities.

YouTube Frames: Additional images were obtained from YouTube videos depicting driving in German towns, adding diversity to the dataset.

Classes:

The dataset includes annotations for the following 18 classes:

- Bicycle

- Building
- Bus
- Car
- Motorcycle
- MyCar
- Person
- Pole
- Rider
- Road
- Sidewalk
- Sky
- Terrain
- Traffic-light
- Traffic-sign
- Train
- Truck
- Vegetation

Annotation Tool

Roboflow: The images were annotated using Roboflow, which facilitated the efficient labeling of objects across 1060 images sourced from a drive link containing multiple folders. You can access the Roboflow project here.

Preprocessing Steps

1. **Resizing**: All images were resized to 640x640 pixels to ensure uniformity and compatibility with the YOLOv8 model.
2. **Splitting**: The dataset was divided into three sets:

   - **Training Set**: 80% of the images.
   - **Validation Set**: 10% of the images.
   - **Test Set**: 10% of the images.

Data Utilization

- **Object Detection**: The YOLOv8 model is trained to detect the 18 specified classes in urban street images.
- **Segmentation**: The detected objects from YOLOv8 are further processed using the Segment Anything Model (SAM) for detailed segmentation.

Conclusion

The dataset used in this project has been carefully selected, annotated, and preprocessed to meet the requirements of object detection and segmentation tasks. By leveraging the Cityscape dataset and additional YouTube Frames, we aim to achieve high accuracy in detecting and segmenting various objects in urban environments. The preprocessing steps, including resizing and data splitting, ensure the dataset is ready for effective training and evaluation of the YOLOv8 and SAM models.