

Name-Akash Rawal

email : akashrawal444@gmail.com

End-to-End Image Instance Segmentation and Tracking Pipeline Using YOLOv11-Seg and Labellerr Platform

1. Assignment Overview

The objective of this assignment was to develop a complete computer vision pipeline for instance segmentation and tracking of vehicles and pedestrians using a **custom dataset**. The workflow included dataset creation, annotation, model training, evaluation, and deployment, **without using pre-trained models**.

Key goals:

- Manual dataset curation and annotation using Labellerr.
- Training YOLOv11-Segmentation from scratch.
- Integration of the trained model into a web application using FastAPI.
- End-to-end workflow validation with hands-on debugging.

2. Dataset and Annotation

- **Dataset Creation:**
 - Total images: **~120 for training, 37 for testing**.
 - Sources: Manually collected from the web and Canva to cover diverse scenarios.
- **Annotation:**
 - **Labellerr platform** was used to perform **instance segmentation** for vehicles and pedestrians.
 - Segmentation masks were created for each relevant object.
 - Annotations were exported in JSON format compatible with YOLOv11-Seg.

3. Model Training and Evaluation

- **Model Used:** YOLO11n-Segmentation.
- **Training:**
 - Trained **from scratch** on the custom dataset (no pre-trained weights).
 - Training duration: **up to 80 epochs** using local CPU and AMD Radeon Graphics.
- **Evaluation:**

- Labellerr SDK was used for testing and validation on **37 test images**, automating prediction uploads and annotation evaluation.

4. Web Application Integration

- **Framework:** FastAPI for interactive deployment.
- **Features:**
 - Users can upload images or videos for segmentation.
 - Returns **segmented outputs** along with JSON annotation files.
- **Output:** Images and annotation data can be used for downstream tasks or further analysis.

5. Bugs and Resolutions

1. During the project, I debugged the YOLOv8 inference outputs and converted them into a COCO-format JSON file. This conversion included extracting image metadata, bounding boxes, segmentation masks, class IDs, and confidence scores, making the dataset compatible for evaluation and further processing
2. **Model Inference Integration:** Initial FastAPI integration issues were debugged to allow smooth processing of images/videos.

6. Key Highlights

- Full end-to-end pipeline with **custom data** and **manual annotation**.
- Training entirely from scratch, emphasizing model adaptation to the dataset.
- Efficient testing and evaluation using Labellerr SDK on test images.
- Interactive web application enables real-time inference for images and videos.

Conclusion:

This project demonstrates a complete instance segmentation and tracking workflow using YOLOv11-Seg, with manual annotation through Labellerr, training from scratch, robust evaluation, and deployment via FastAPI. The workflow emphasizes **hands-on expertise in dataset preparation, annotation, model training**.