YOLO Roadscene Detection Project Report

Abstract

This project applies YOLOv8 for detecting multiple object categories in road scenes, including road, vehicle, pedestrian, traffic sign, building, sky, and other. The dataset was prepared in YOLO format, trained in Google Colab, and evaluated using precision, recall, and mAP. We also provide CLI and Gradio demos for inference.

1. Introduction

Real-time road-scene understanding is vital for autonomous driving and traffic analytics. YOLOv8 offers fast inference and competitive accuracy, making it suitable for deployment.
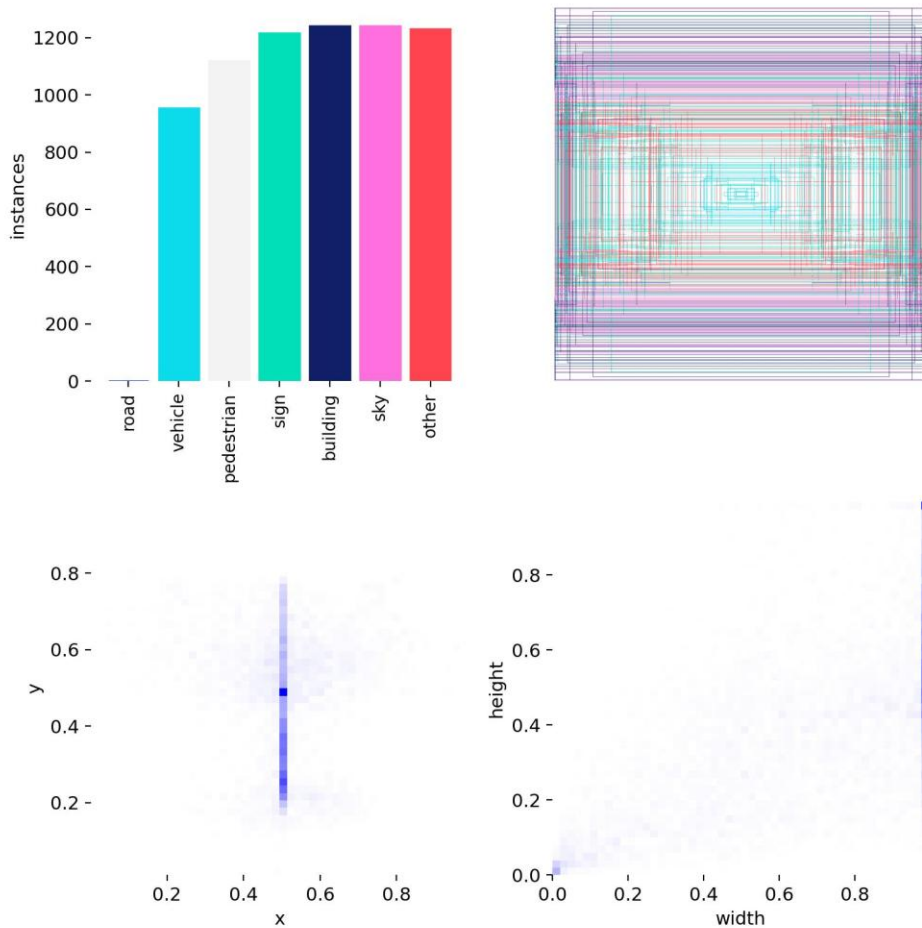
2. Dataset

Dataset:

India Driving Dataset (IDD), converted to YOLO format.
Classes: road, vehicle, pedestrian, sign, building, sky, other.
Split example: ~1242 training and 138 validation images.
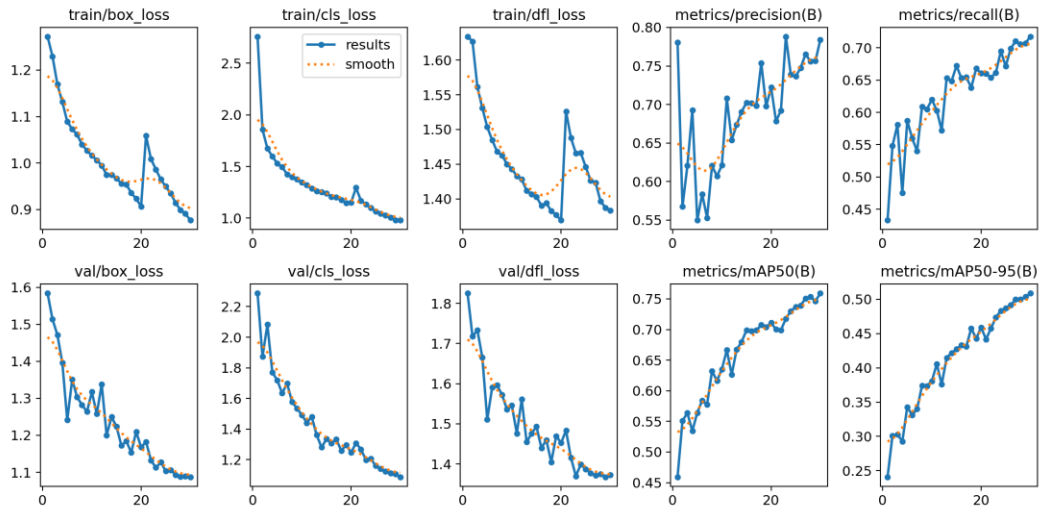
Class distribution:

3. Methodology

- Data preparation and YOLO formatting
- Model: YOLOv8n pre-trained weights
- Training: ~30 epochs, batch=16, imgsz=640 (Colab)
- Validation: Precision, Recall, mAP@50, mAP@50-95
- Export: ONNX for portability

4. Results

Final metrics (last epoch):
• Precision: 0.784
• Recall: 0.718
• mAP@50: 0.759
• mAP@50-95: 0.509

Training & Validation Curves:

Confusion Matrix:

Sample Predictions

5. Deployment

- CLI: infer.py (run predictions on images/videos)
- Web: gradio_app.py (browser demo)
- Model export: best.onnx for integration

6. Conclusion & Future Work

YOLOv8 successfully detects road scene objects with strong accuracy. Future work: use larger variants (YOLOv8m/l), multi-dataset training, and edge deployment.