

YOLO Model Testing Report

Model:	yolov8s
Model Size:	21.5 MB
Parameters:	11.2 M
FLOPs (640x640):	28.82 GFLOPs
Run Name:	yolov8s_testing_20251127_211340
Timestamp:	2025-11-27 21:25:35
Dataset:	bdd100k_yolo_limited - test split
Images Processed:	20000
Total Execution Time	101.54s
IoU Threshold:	0.5

Inference Performance

Metric	Value
Average Inference Time	1.77 ms per image
FPS (Frames Per Second)	564.55

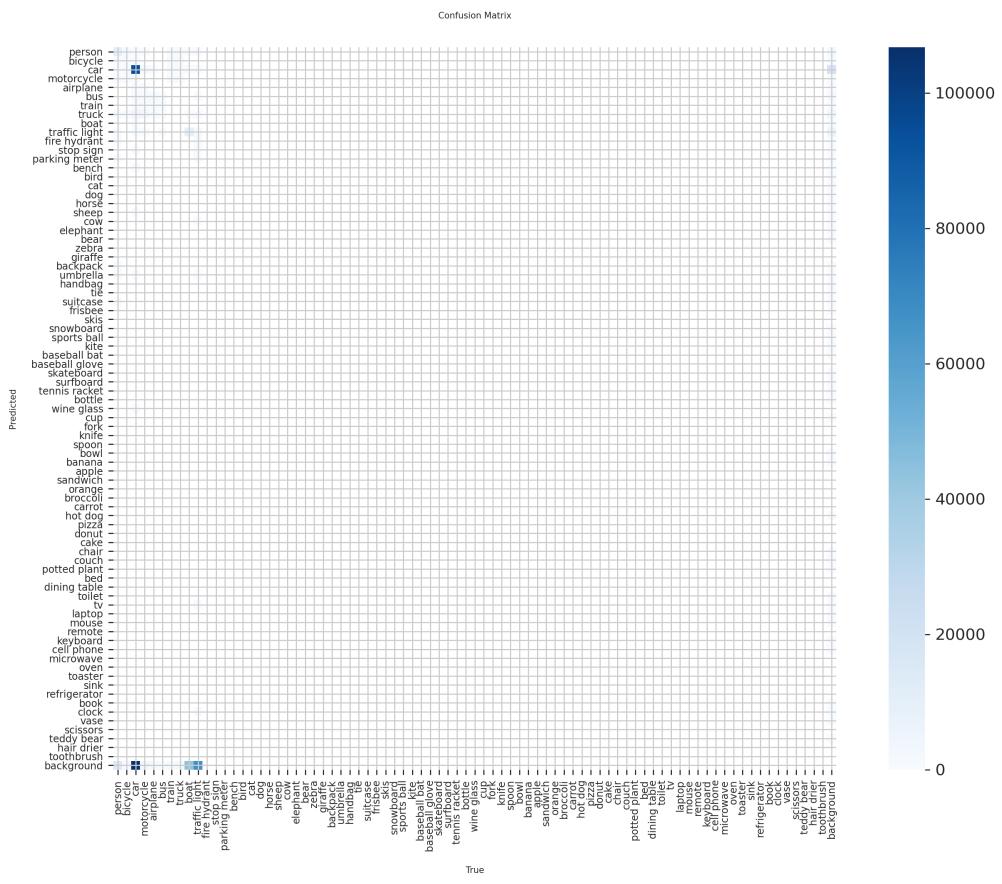
Note: Inference time includes preprocessing, model inference, and postprocessing.

Overall Accuracy Metrics

Metric	Value
Precision	0.2847
Recall	0.6668
F1-Score	0.3991
mAP@0.5	0.1098
mAP@0.5:0.95	0.0592
True Positives	104706
False Positives	263021
False Negatives	52318

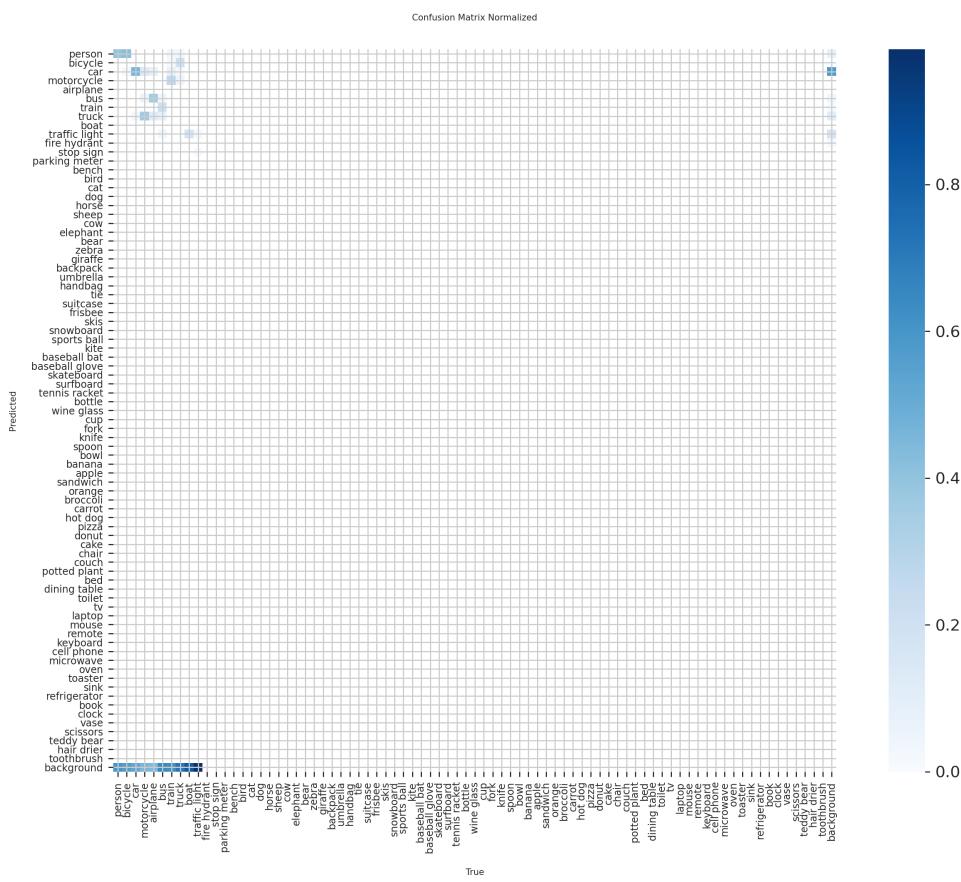
Confusion Matrices

Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

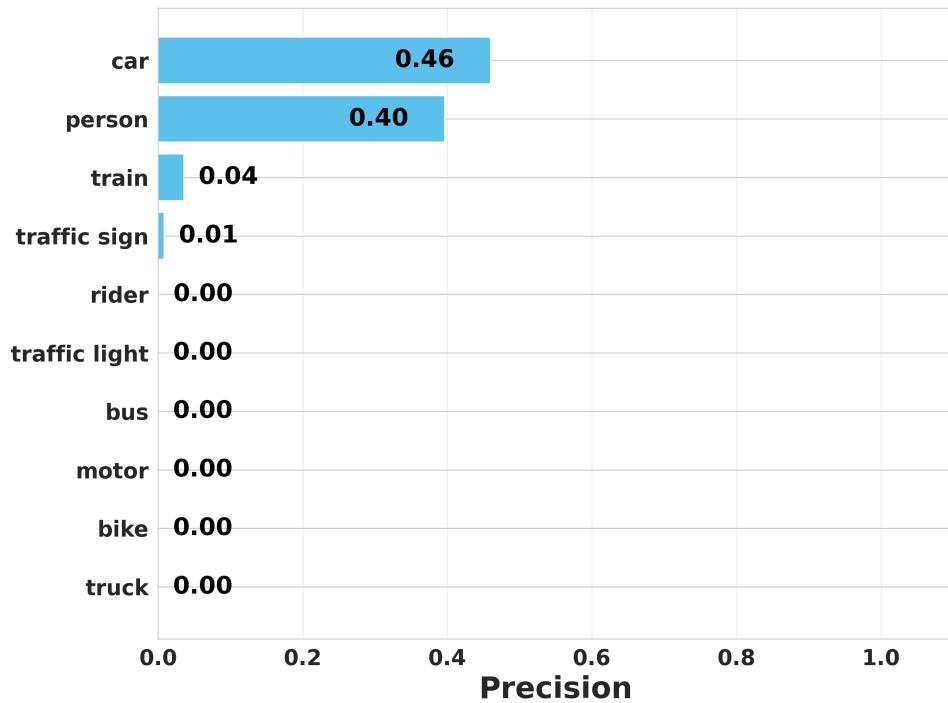
Confusion Matrix (Normalized)



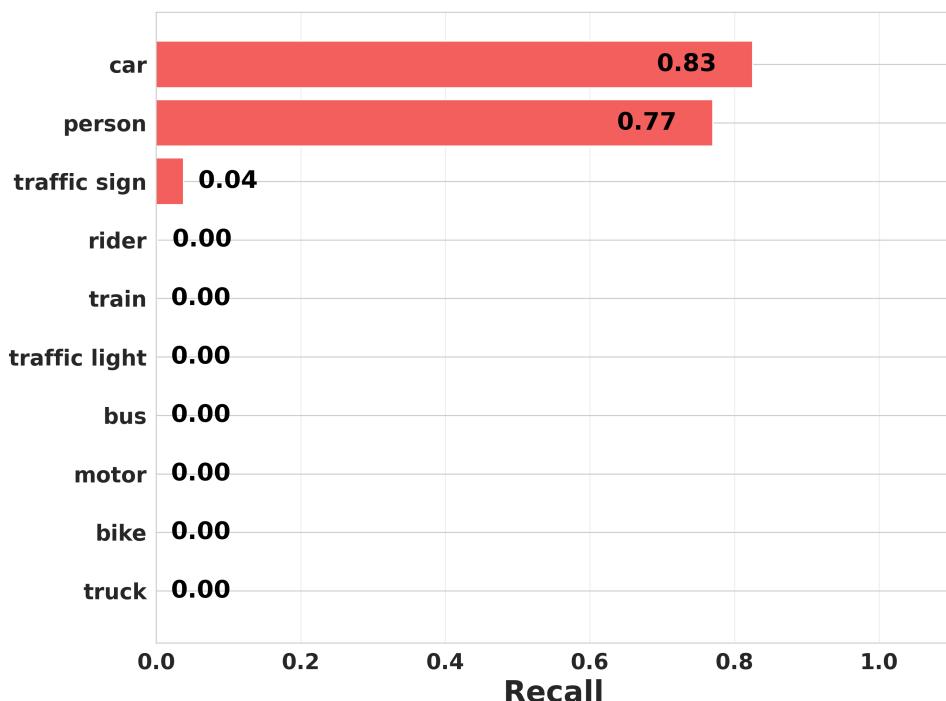
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

Performance Visualizations

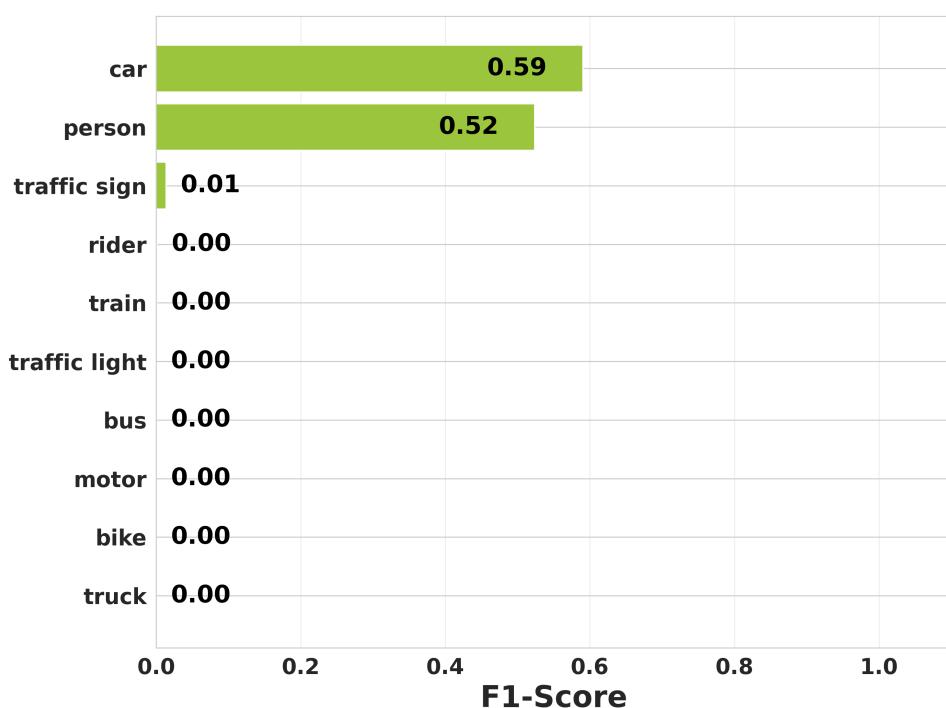
Precision by Class



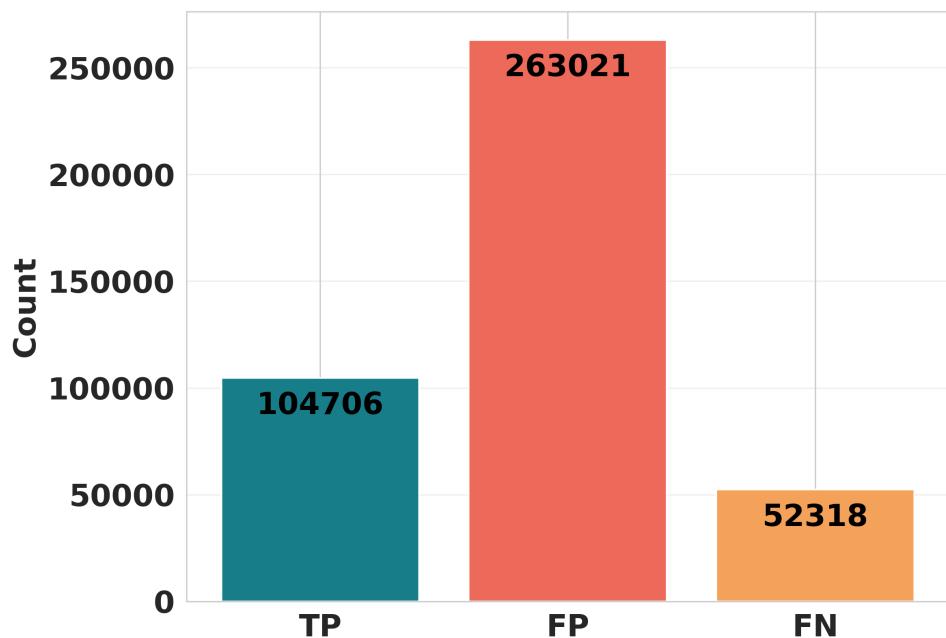
Recall by Class



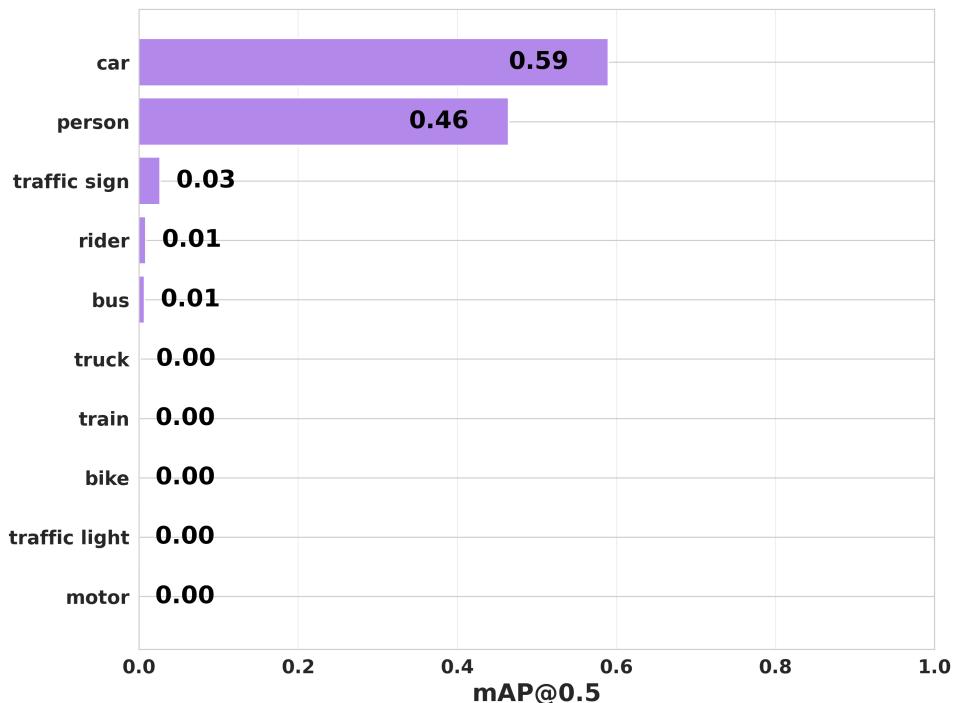
F1-Score by Class



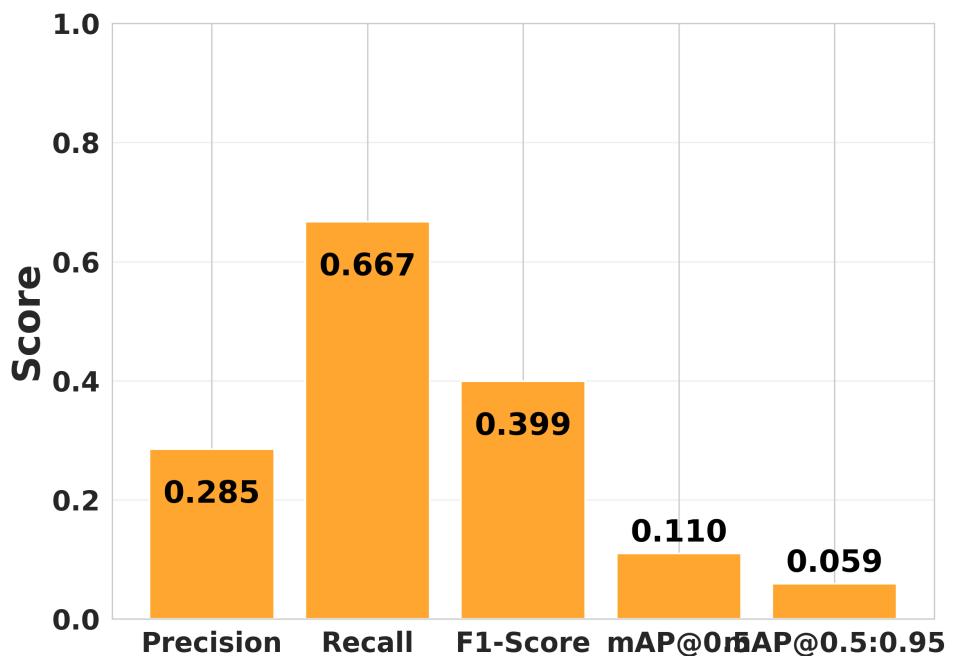
Overall Detection Outcomes



mAP@0.5 by Class



Overall Metrics



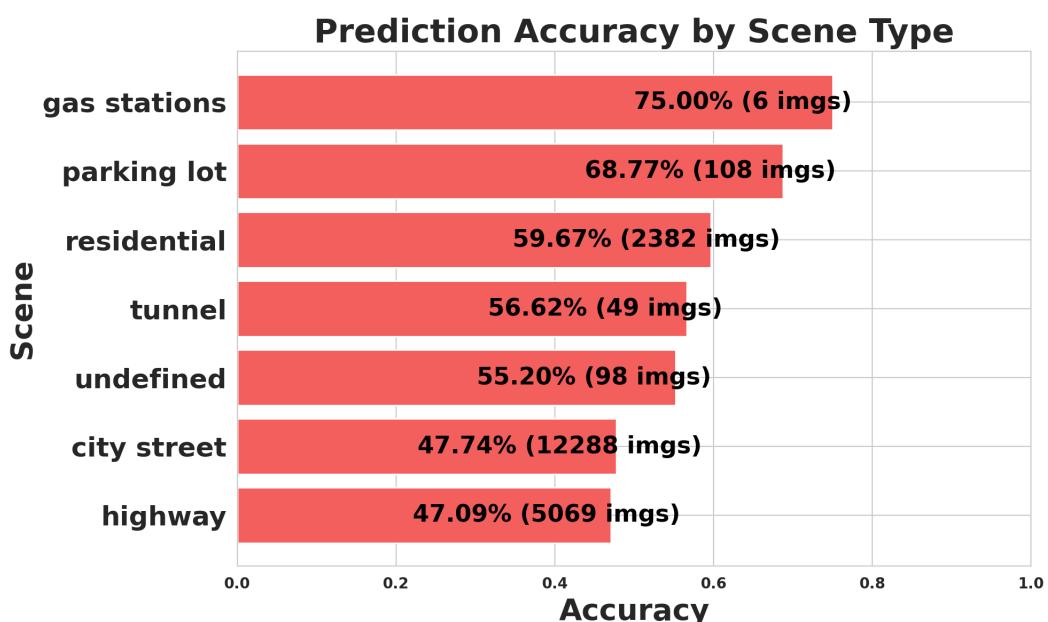
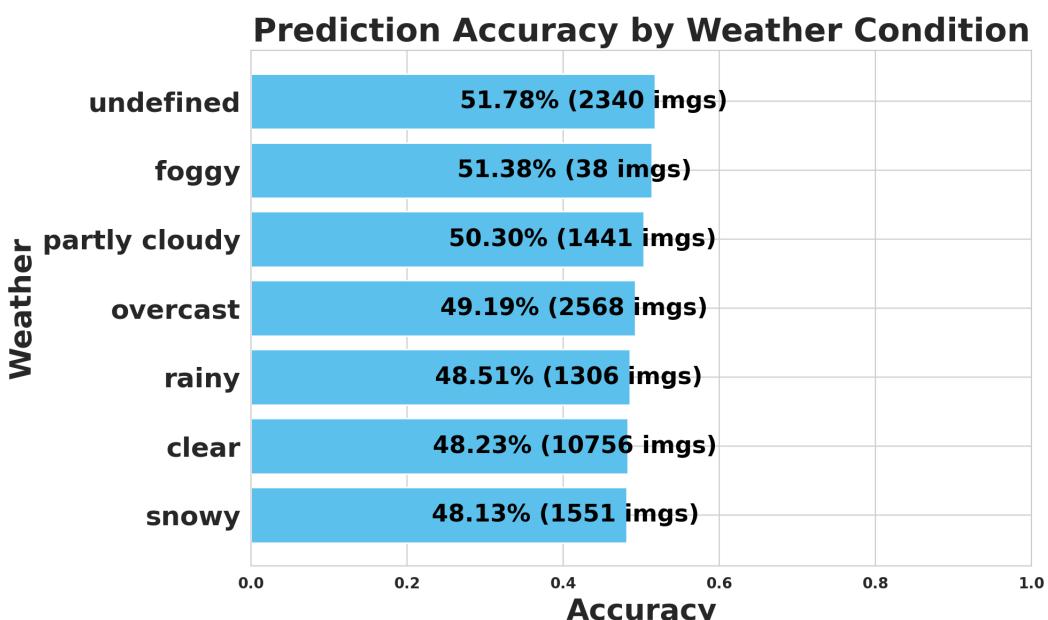
Per-Class Performance

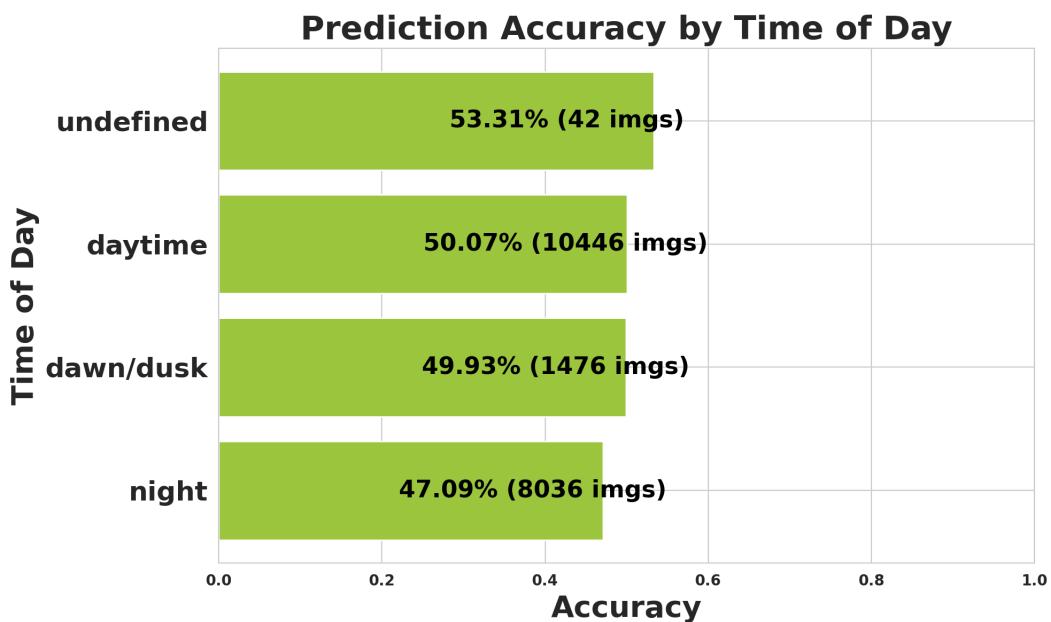
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	9776	14874	2920	0.3966	0.7700	0.5235	0.4645
rider	1	1293	567	0.0008	0.0018	0.0011	0.0085
car	94355	110794	20000	0.4599	0.8251	0.5906	0.5897
truck	0	8704	466	0.0000	0.0000	0.0000	0.0013
bus	0	3217	26	0.0000	0.0000	0.0000	0.0069
train	1	27	3345	0.0357	0.0003	0.0006	0.0002
motor	0	841	224	0.0000	0.0000	0.0000	0.0001
bike	0	1998	10216	0.0000	0.0000	0.0000	0.0002
traffic light	0	52840	12	0.0000	0.0000	0.0000	0.0001
traffic sign	573	68433	14542	0.0083	0.0379	0.0136	0.0263

Comprehensive Failure Analysis

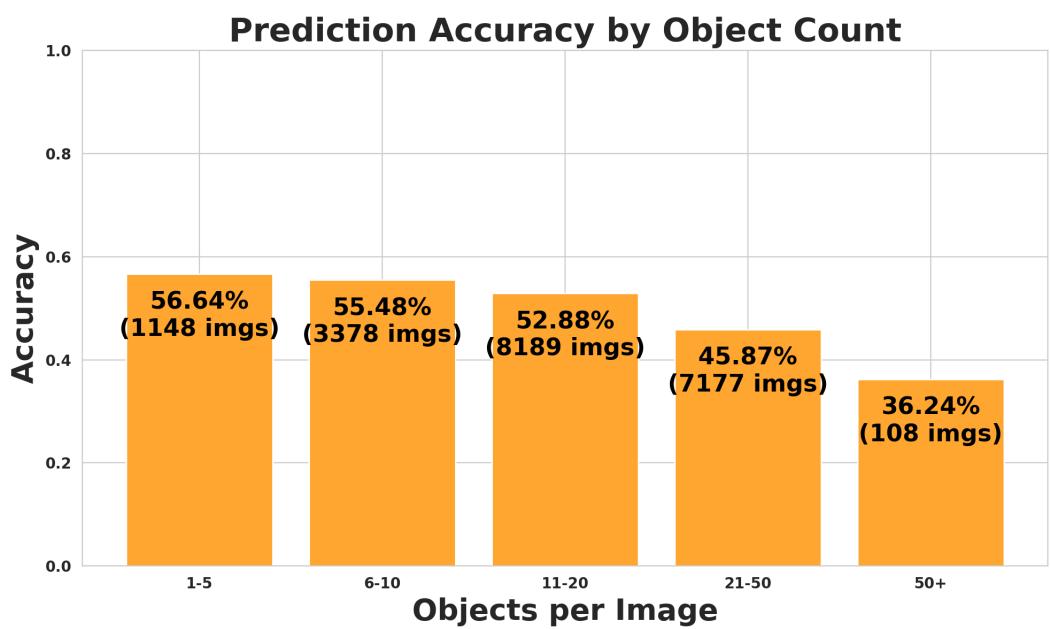
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	180177
Overall Accuracy	49.00%

Accuracy by Environmental Attributes

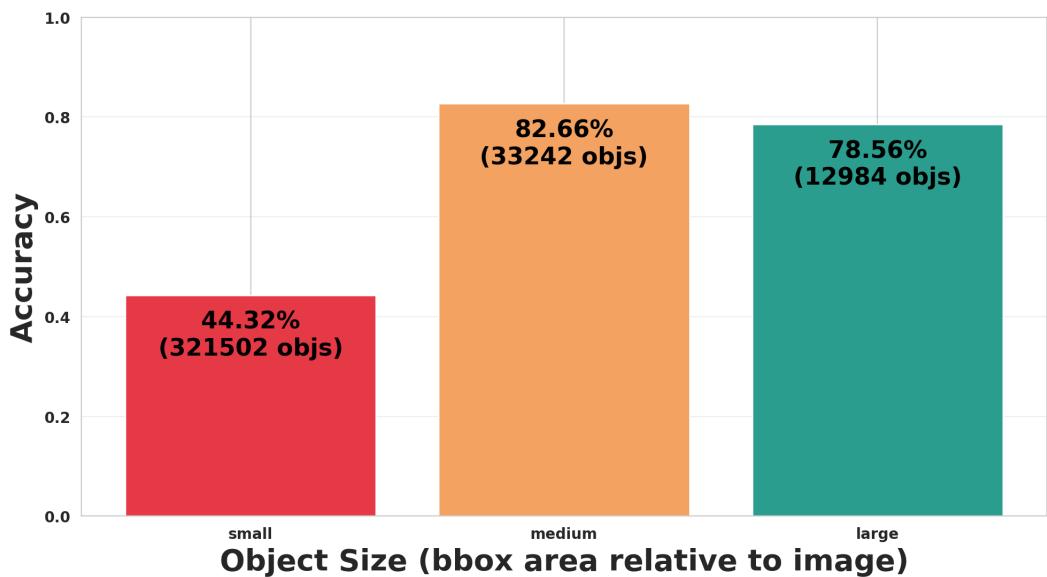




Accuracy by Object Characteristics

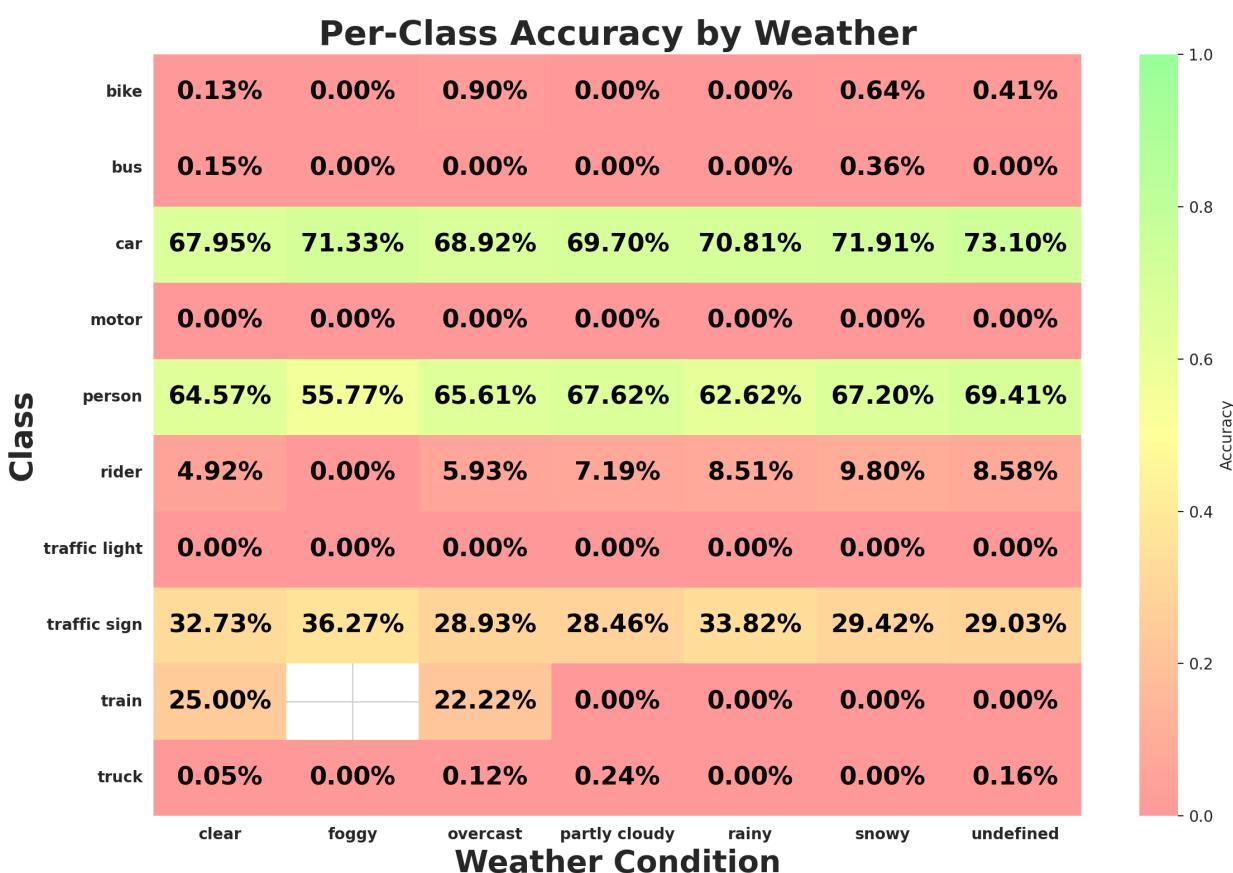
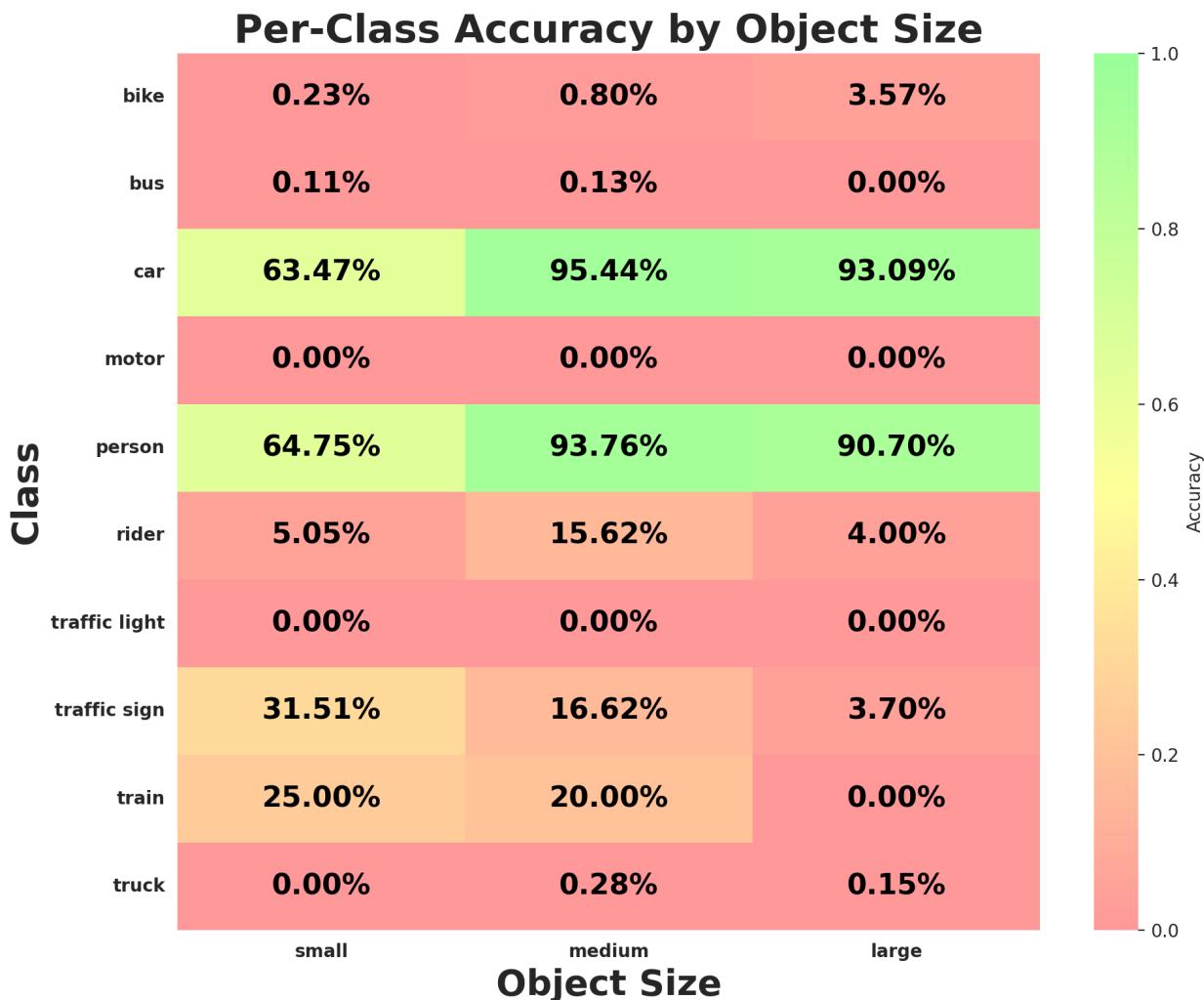


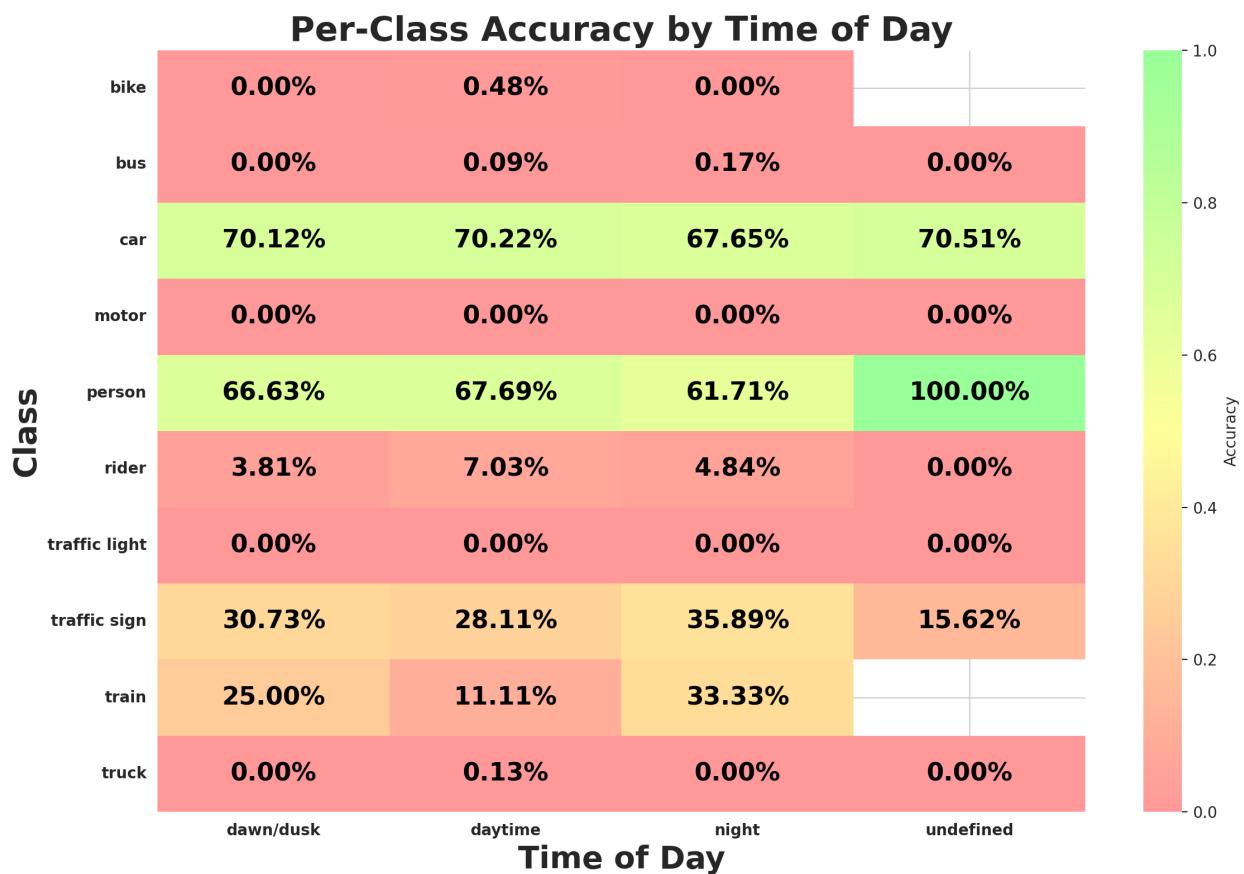
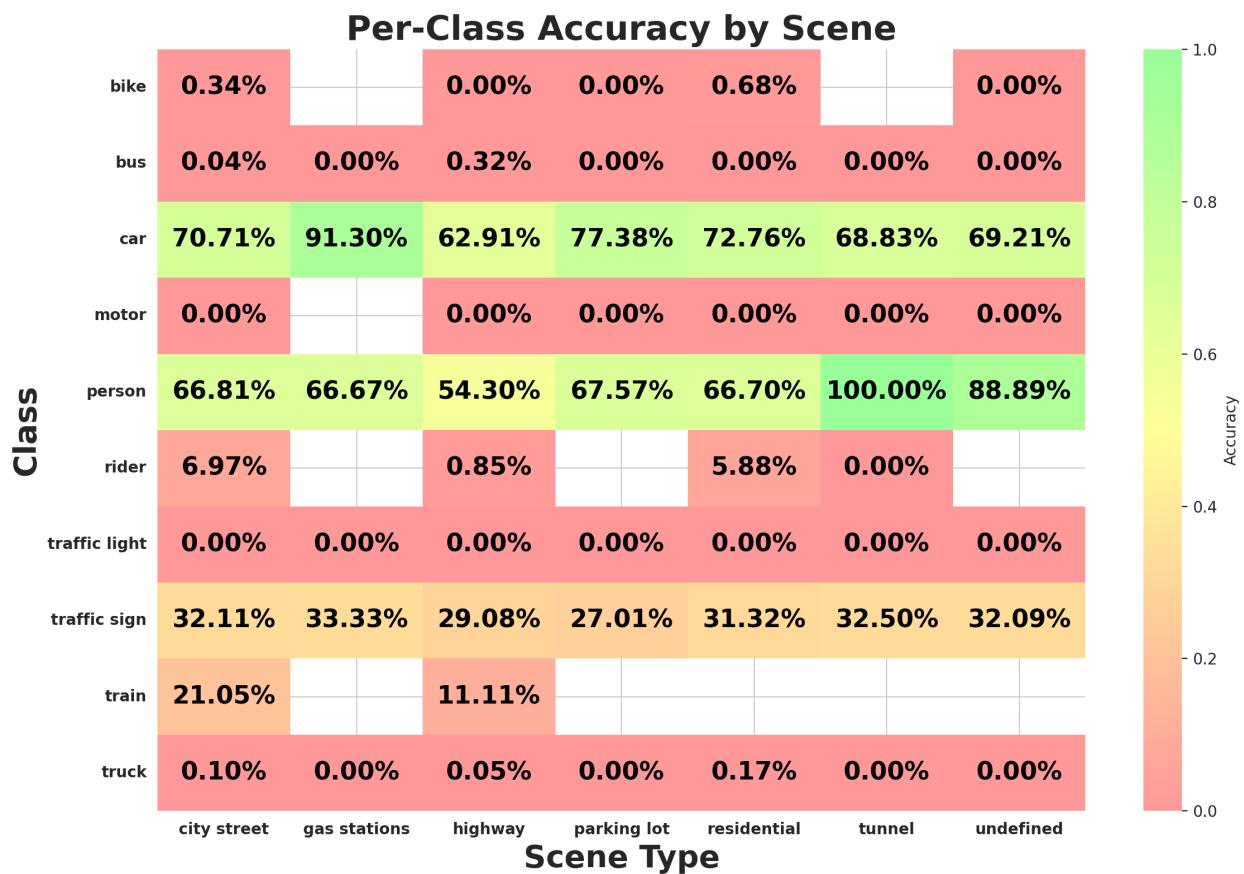
Prediction Accuracy by Object Size



Note: Object size categories are defined based on bounding box area as a percentage of total image area: **Small** objects occupy less than 1% of the image (typically distant or small objects), **Medium** objects occupy between 1% and 5% of the image (moderate-sized objects at medium distances), and **Large** objects occupy more than 5% of the image (close or large objects). These thresholds help identify how object scale affects detection accuracy.

Per-Class Performance Analysis

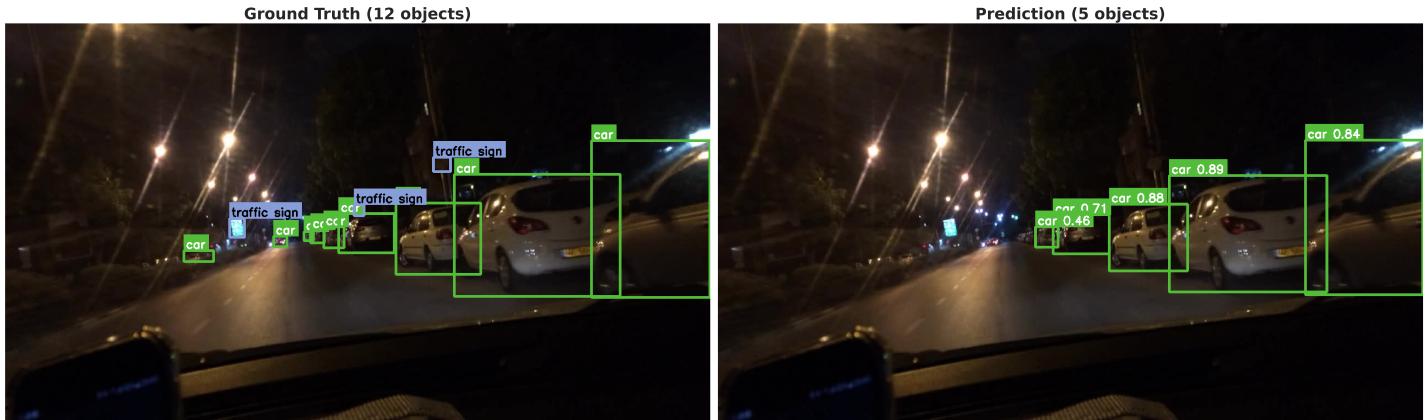




Sample Predictions: Ground Truth vs Model

Sample #1: e8a5ac01-970d1b53.jpg

Ground Truth Objects: 12 | Predicted Objects: 5
Weather: rainy | Scene: highway | Time of Day: night



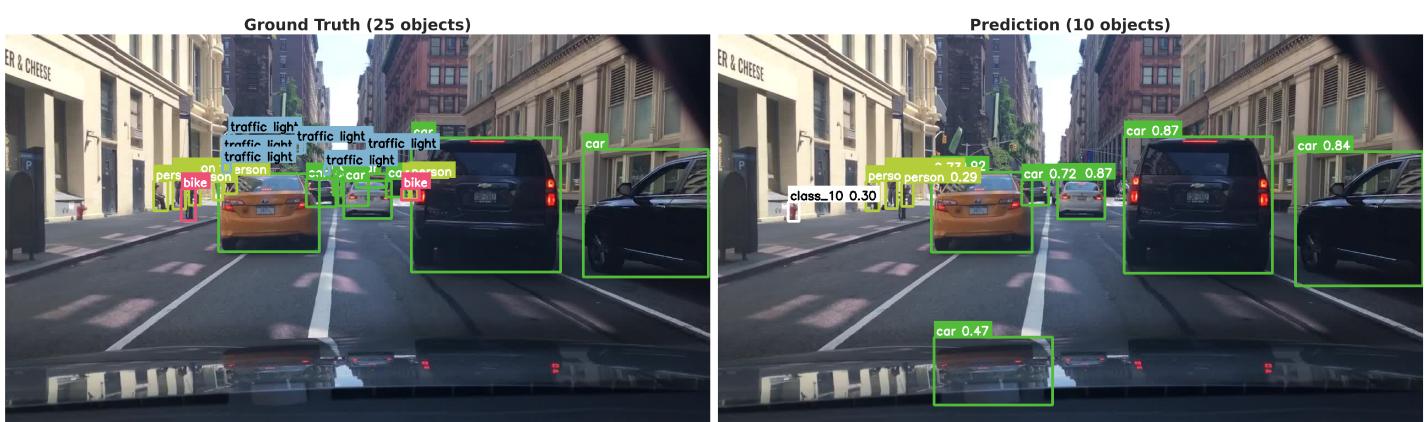
Sample #2: cec118b2-bc96966d.jpg

Ground Truth Objects: 4 | Predicted Objects: 6
Weather: undefined | Scene: highway | Time of Day: daytime



Sample #3: cecda80e-b32ec9ca.jpg

Ground Truth Objects: 25 | Predicted Objects: 10
Weather: clear | Scene: city street | Time of Day: daytime



Sample #4: e88fcfb8-a0ff727c.jpg

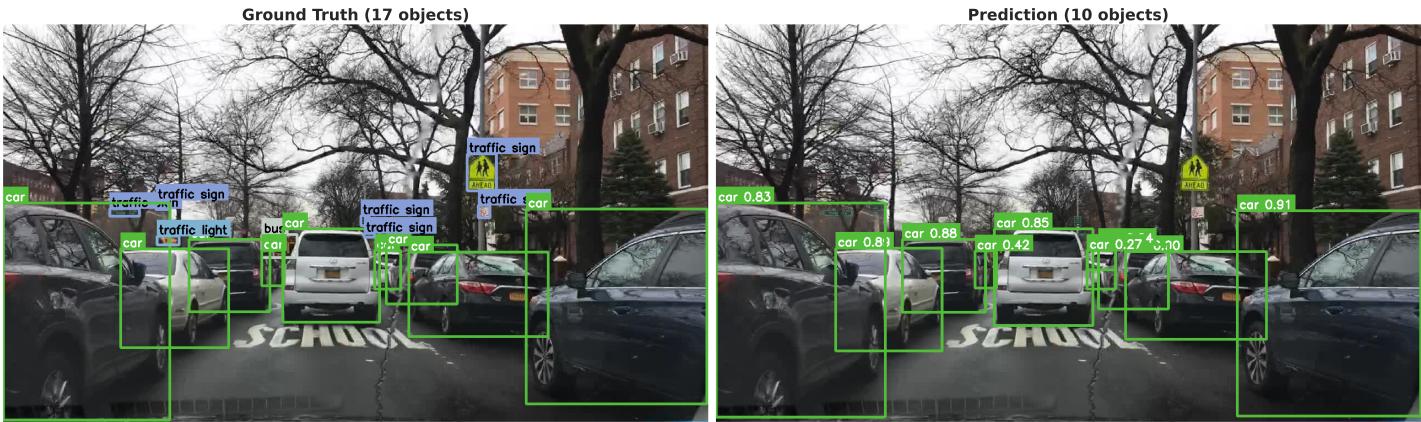
Ground Truth Objects: 15 | Predicted Objects: 6
Weather: clear | Scene: highway | Time of Day: daytime



Sample #5: db9fa2d7-42c50b21.jpg

Ground Truth Objects: 17 | Predicted Objects: 10

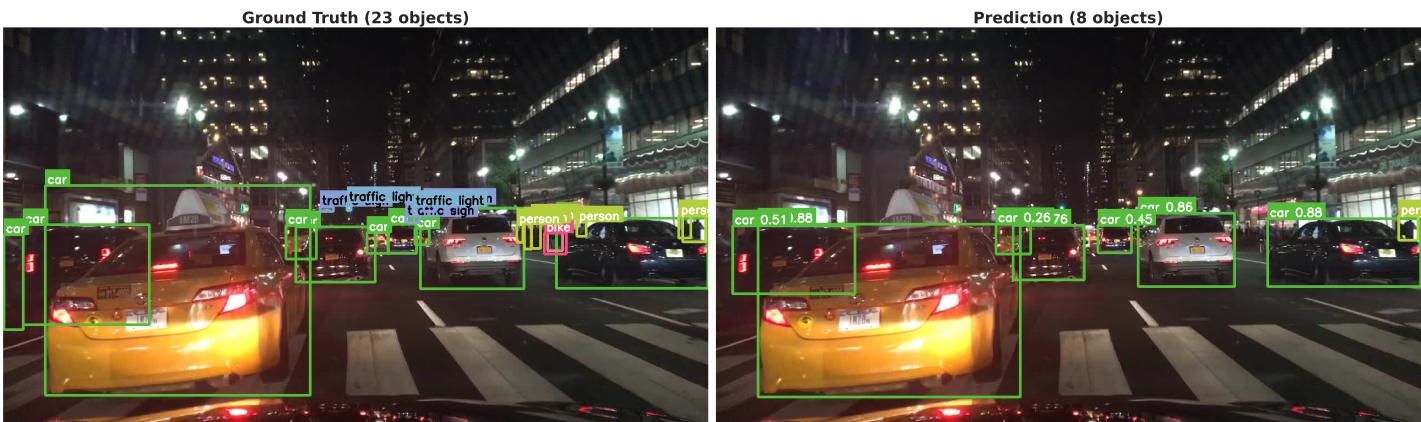
Weather: overcast | Scene: residential | Time of Day: daytime



Sample #6: d665ad15-7042cccd1.jpg

Ground Truth Objects: 23 | Predicted Objects: 8

Weather: clear | Scene: city street | Time of Day: night



Generated by YOLO Testing Pipeline