

YOLO Model Testing Report

Model:	yolov10s
Model Size:	15.9 MB
Parameters:	8.1 M
FLOPs (640x640):	25.11 GFLOPs
Run Name:	yolov10s_testing_20251127_232138
Timestamp:	2025-11-27 23:37:34
Dataset:	bdd100k_yolo_limited - test split
Images Processed:	20000
Total Execution Time	95.07s
IoU Threshold:	0.5

Inference Performance

Metric	Value
Average Inference Time	1.26 ms per image
FPS (Frames Per Second)	793.72

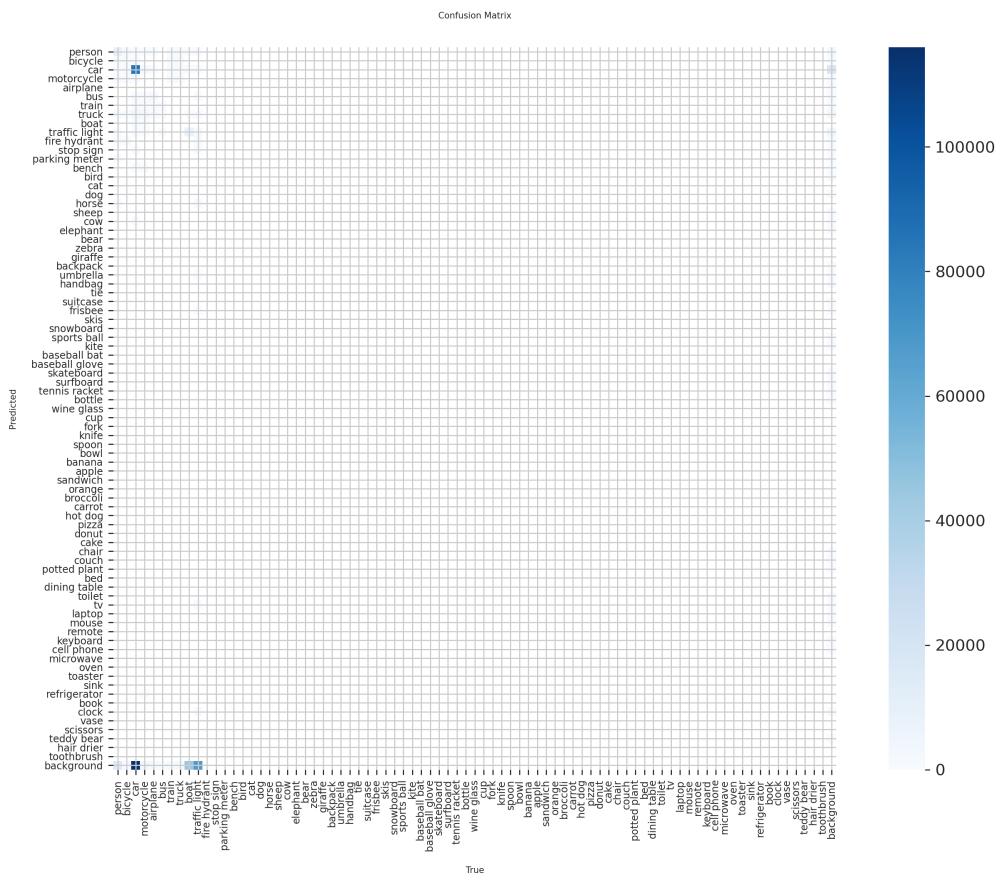
Note: Inference time includes preprocessing, model inference, and postprocessing.

Overall Accuracy Metrics

Metric	Value
Precision	0.2575
Recall	0.6717
F1-Score	0.3723
mAP@0.5	0.1059
mAP@0.5:0.95	0.0582
True Positives	94691
False Positives	273036
False Negatives	46272

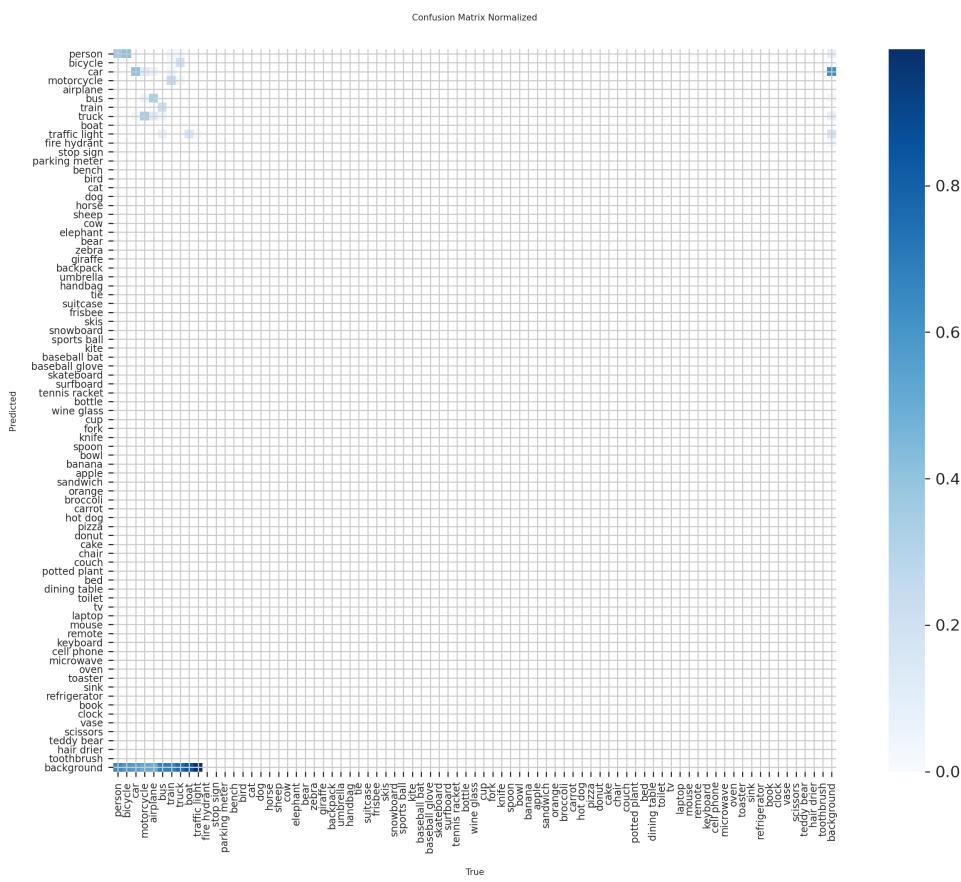
Confusion Matrices

Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

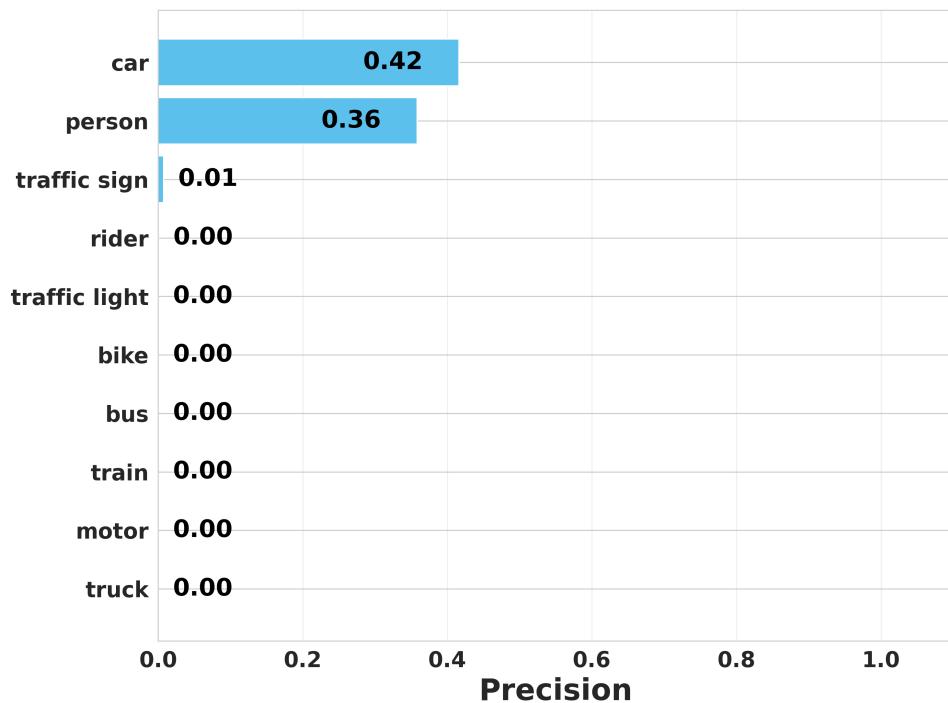
Confusion Matrix (Normalized)



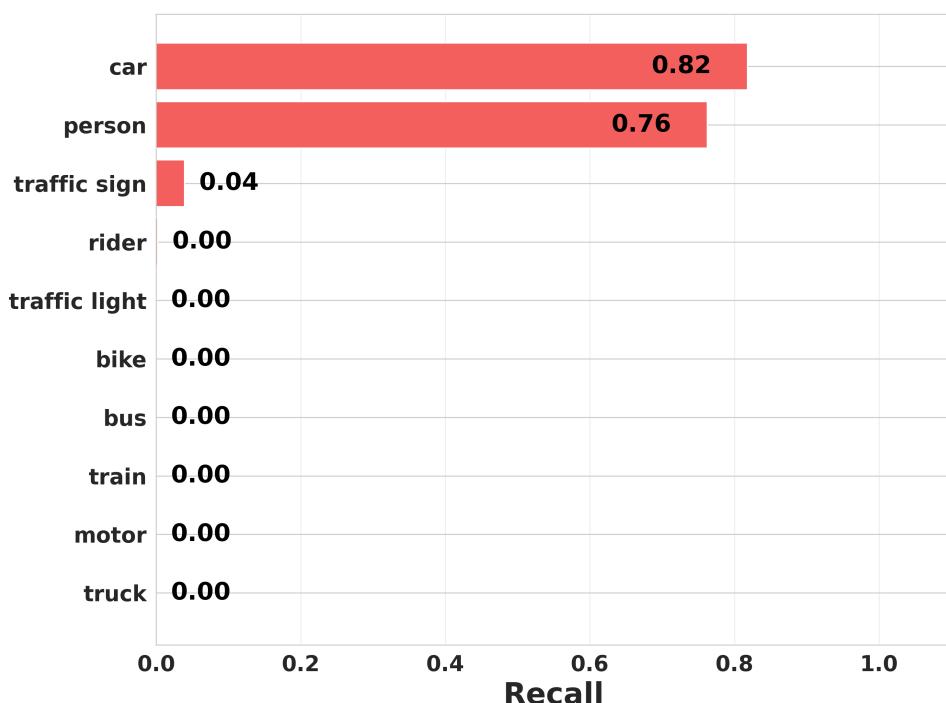
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

Performance Visualizations

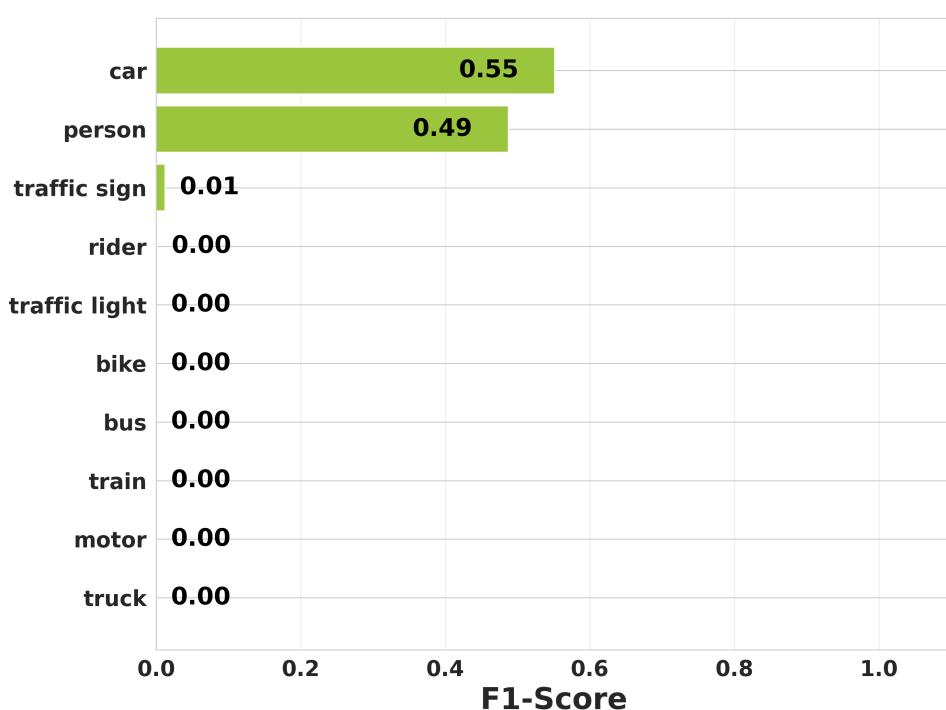
Precision by Class



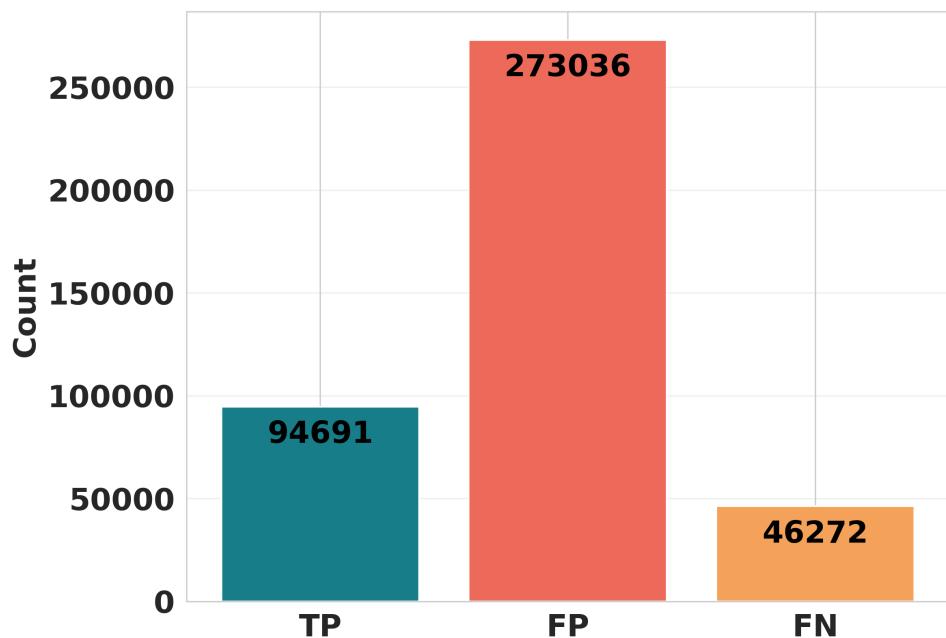
Recall by Class



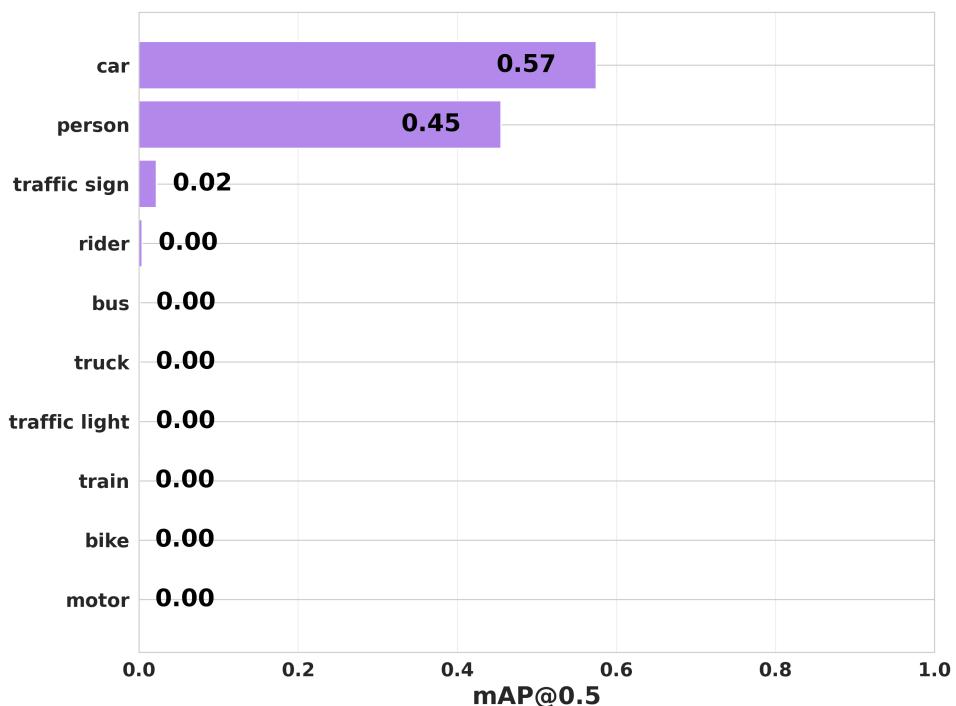
F1-Score by Class



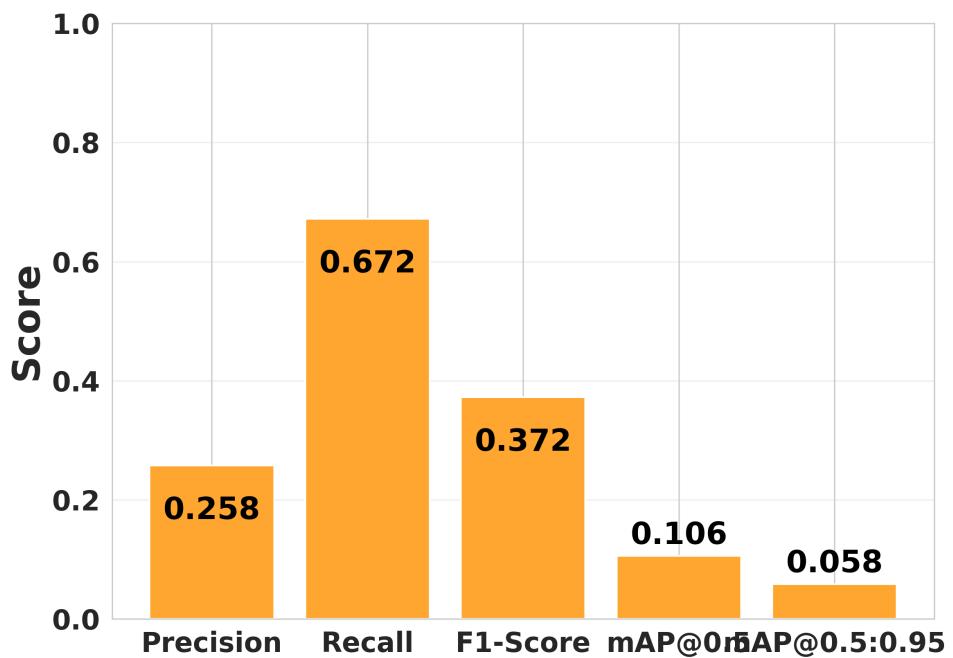
Overall Detection Outcomes



mAP@0.5 by Class



Overall Metrics



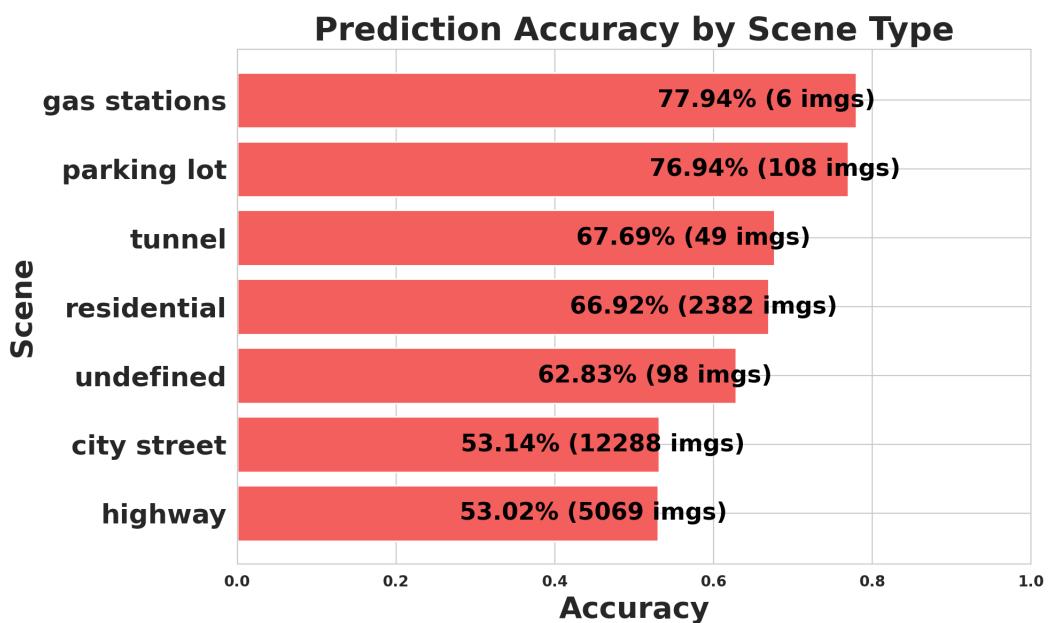
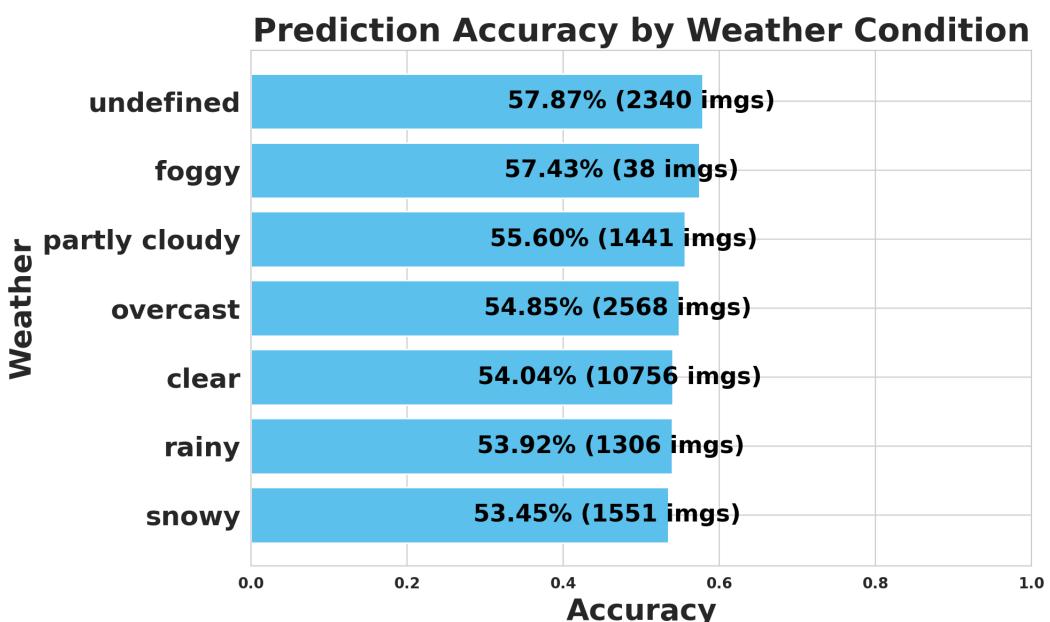
Per-Class Performance

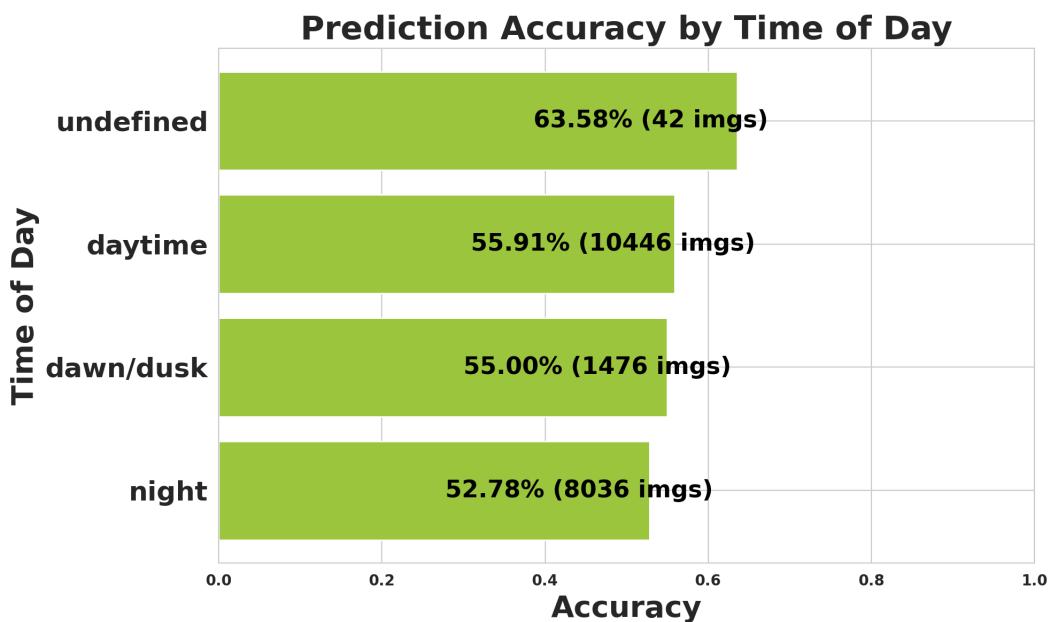
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	8831	15819	2747	0.3583	0.7627	0.4875	0.4549
rider	1	1293	508	0.0008	0.0020	0.0011	0.0041
car	85350	119799	18987	0.4160	0.8180	0.5516	0.5746
truck	0	8704	401	0.0000	0.0000	0.0000	0.0009
bus	0	3217	6	0.0000	0.0000	0.0000	0.0013
train	0	28	2345	0.0000	0.0000	0.0000	0.0001
motor	0	841	59	0.0000	0.0000	0.0000	0.0000
bike	0	1998	8723	0.0000	0.0000	0.0000	0.0001
traffic light	0	52840	2	0.0000	0.0000	0.0000	0.0007
traffic sign	509	68497	12494	0.0074	0.0391	0.0124	0.0218

Comprehensive Failure Analysis

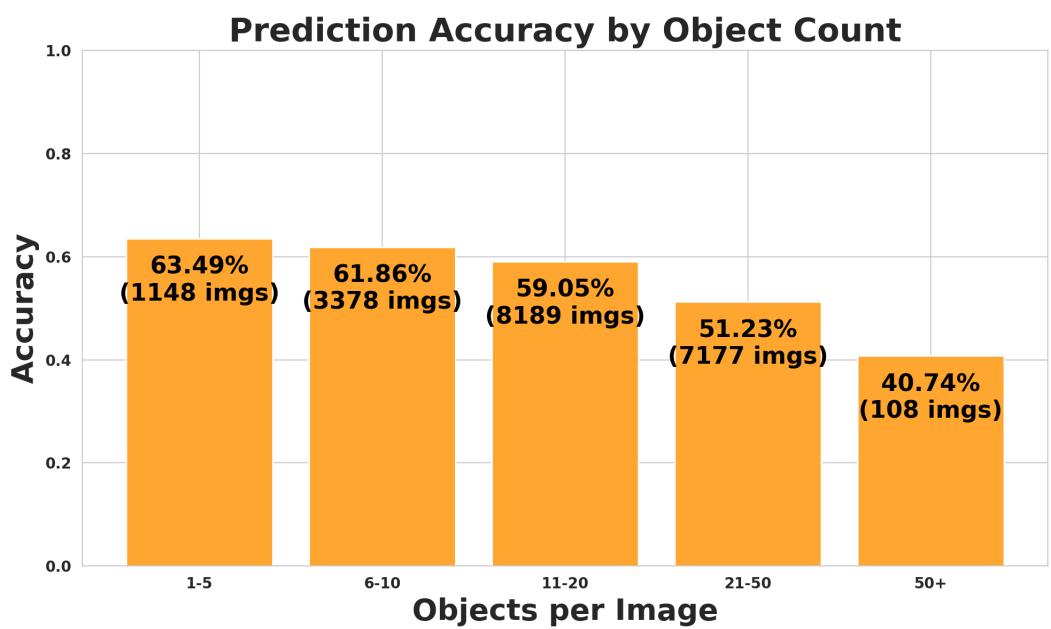
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	201233
Overall Accuracy	54.72%

Accuracy by Environmental Attributes

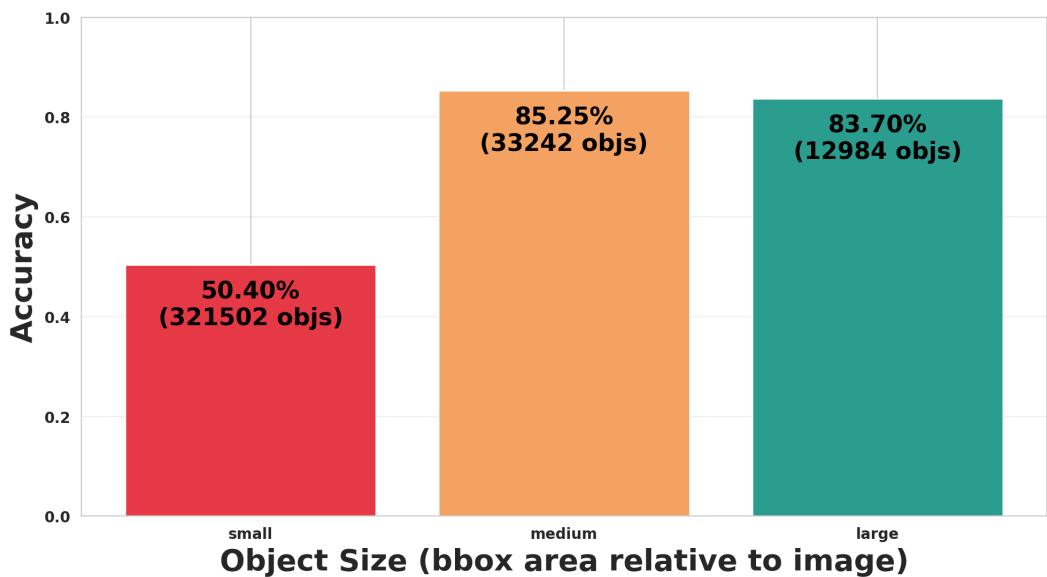




Accuracy by Object Characteristics

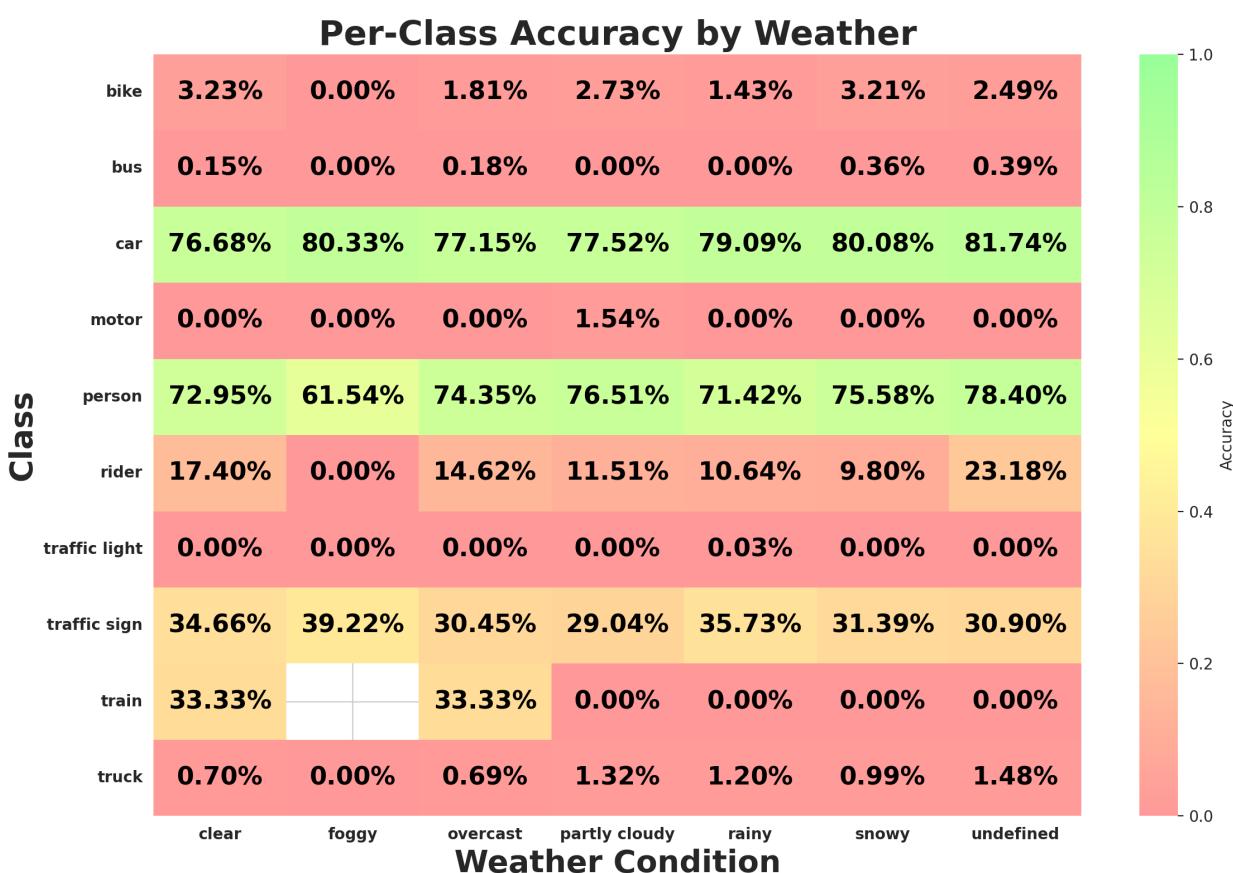
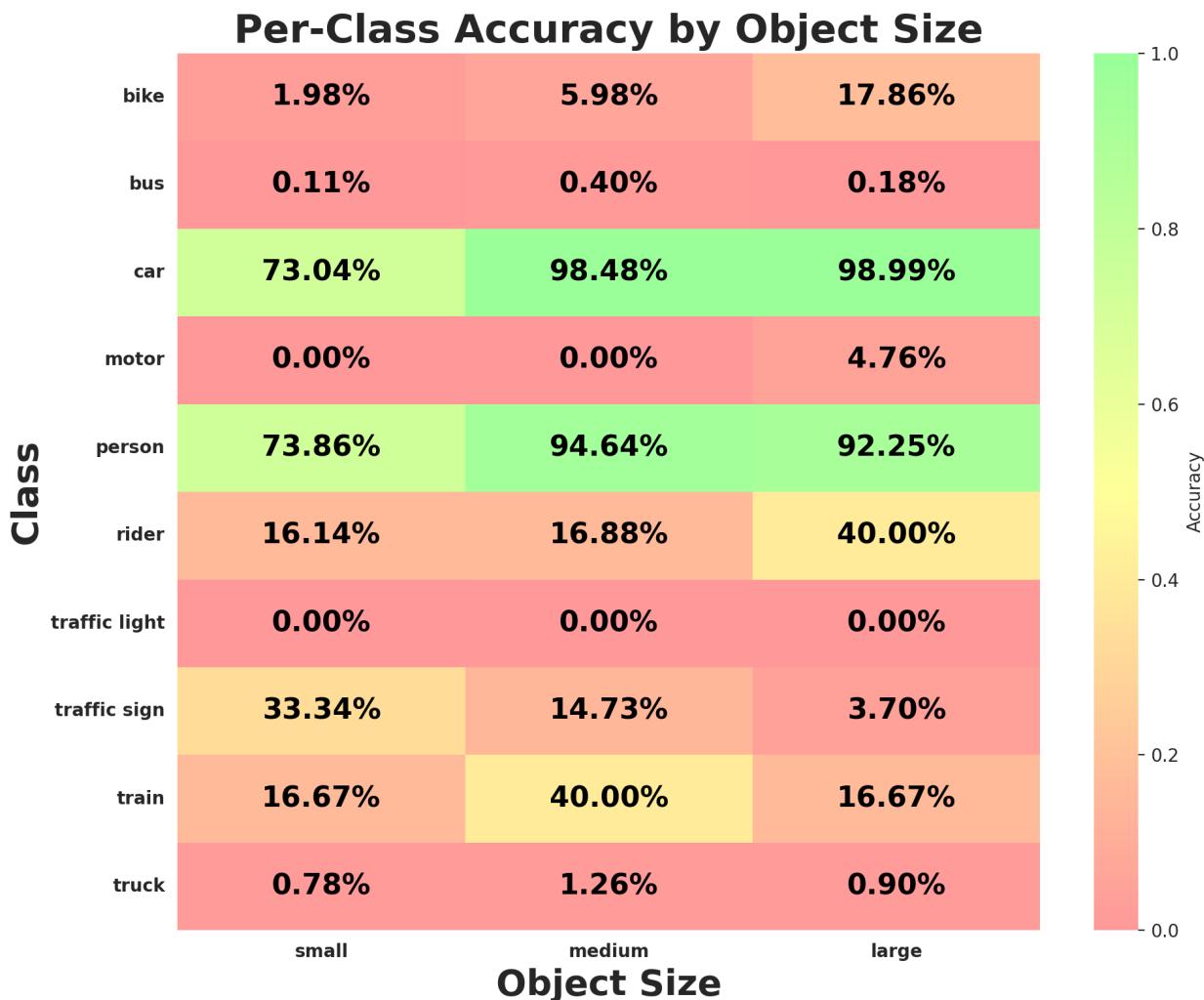


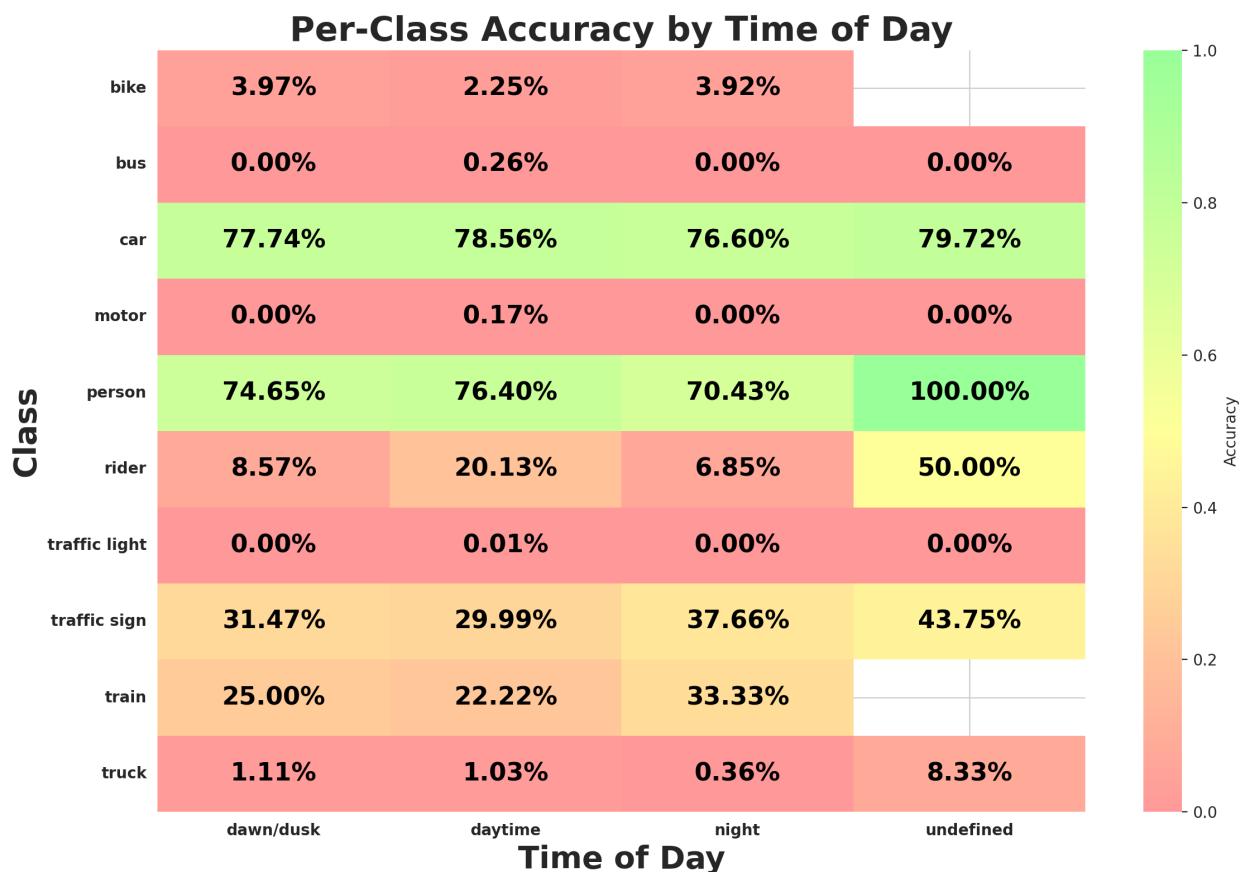
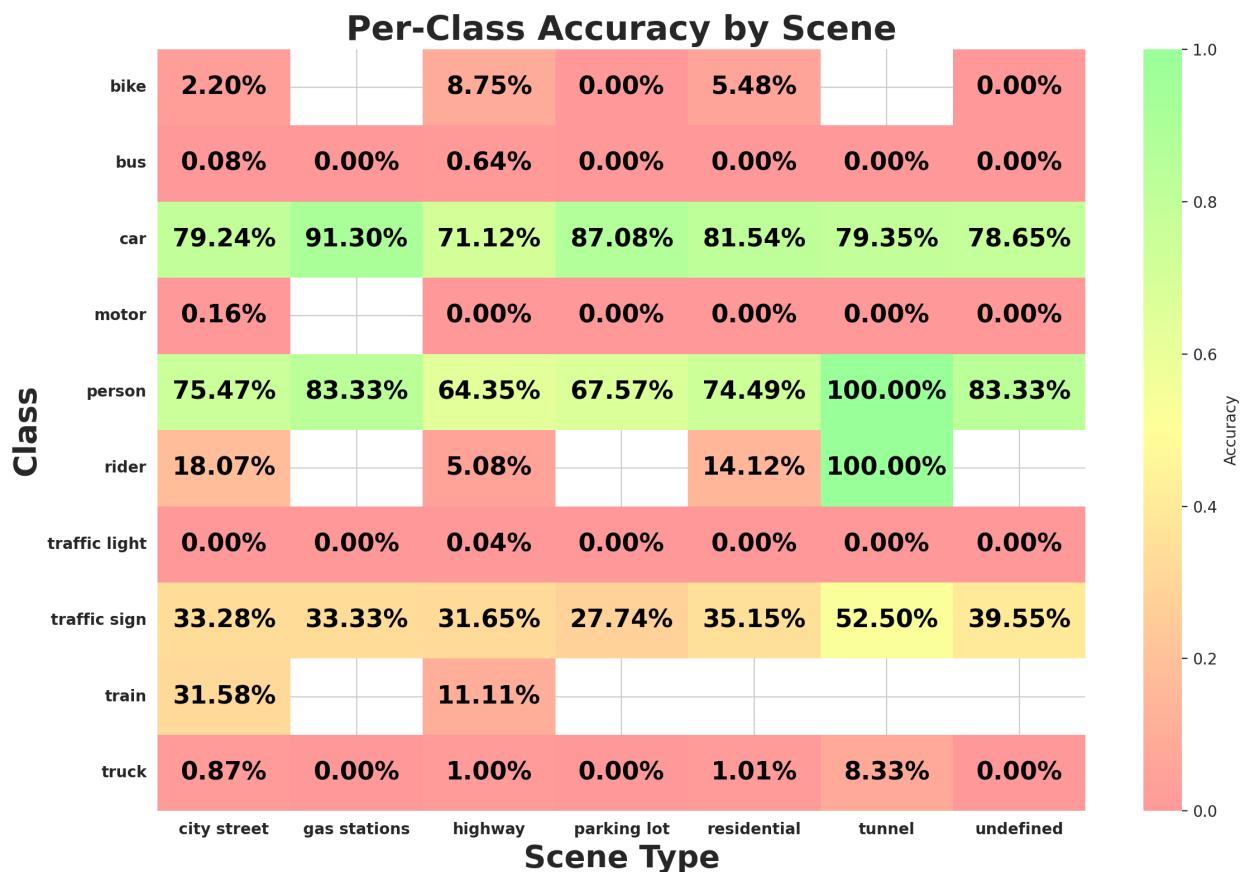
Prediction Accuracy by Object Size



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

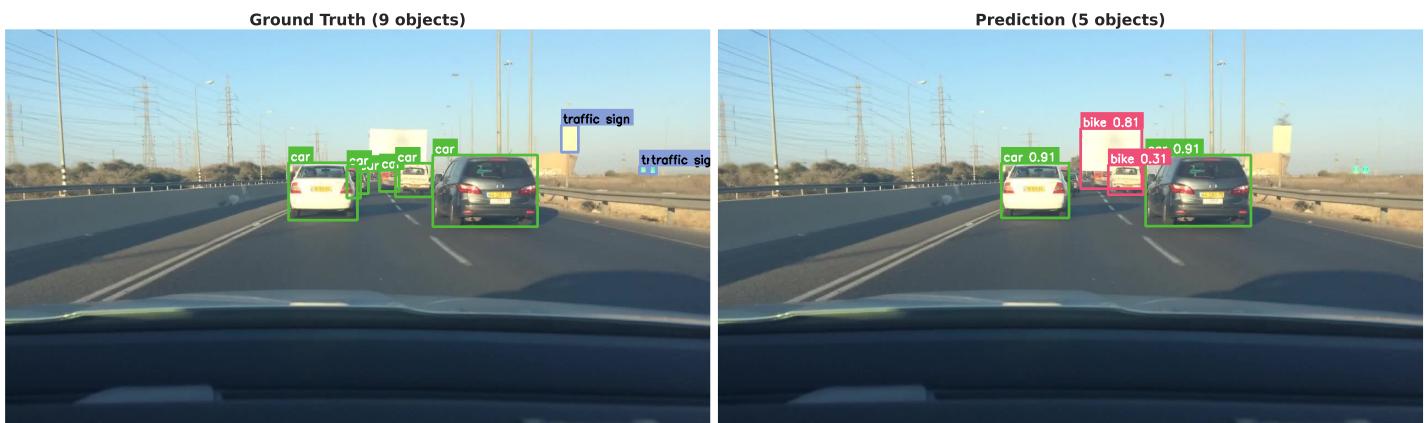
Sample #1: df06cde1-889ab87e.jpg

Ground Truth Objects: 17 | Predicted Objects: 10
Weather: clear | Scene: city street | Time of Day: night



Sample #2: ea9d5697-de24ae59.jpg

Ground Truth Objects: 9 | Predicted Objects: 5
Weather: clear | Scene: highway | Time of Day: daytime



Sample #3: f8284e36-08ff9271.jpg

Ground Truth Objects: 24 | Predicted Objects: 8
Weather: rainy | Scene: highway | Time of Day: daytime



Sample #4: dff0bbbf-e34fed77.jpg

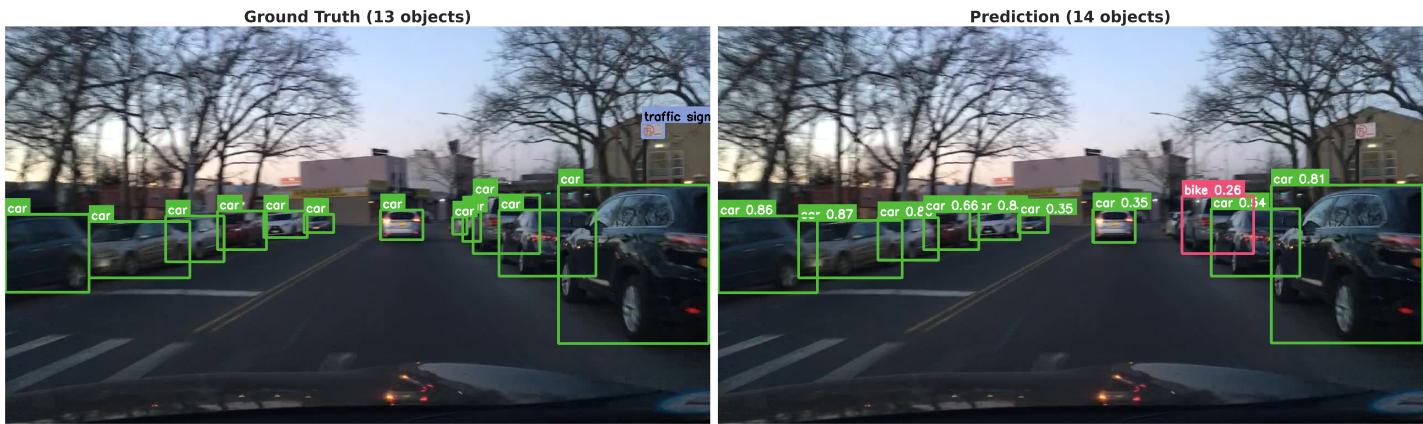
Ground Truth Objects: 7 | Predicted Objects: 6
Weather: clear | Scene: city street | Time of Day: night



Sample #5: e02ea06e-58642176.jpg

Ground Truth Objects: 13 | Predicted Objects: 14

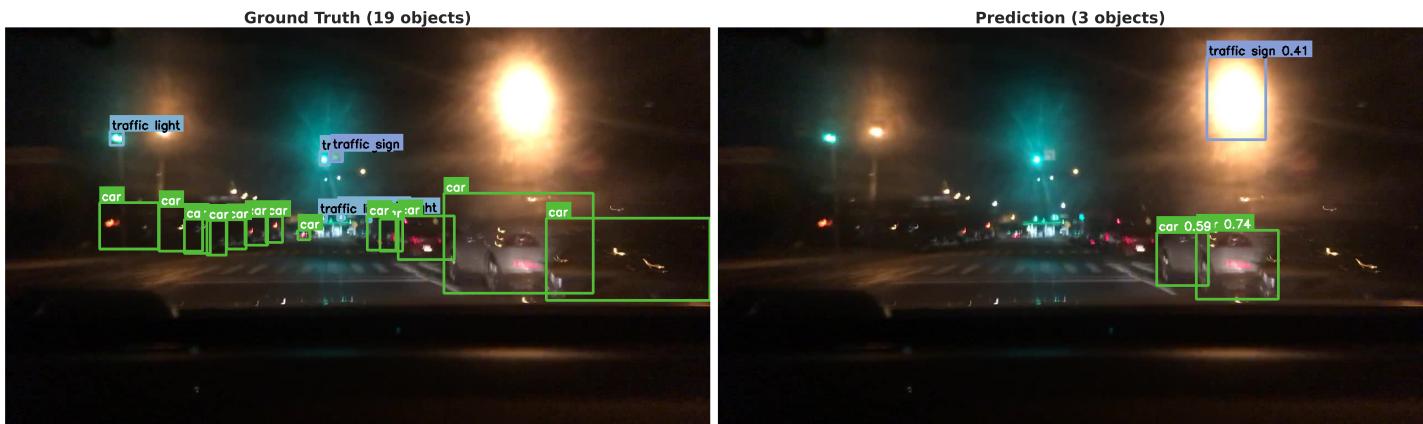
Weather: clear | Scene: city street | Time of Day: dawn/dusk



Sample #6: f05132ee-ca6fb293.jpg

Ground Truth Objects: 19 | Predicted Objects: 3

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



Generated by YOLO Testing Pipeline