

YOLO Model Testing Report

Model:	yolov8m
Model Size:	49.7 MB
Parameters:	25.9 M
FLOPs (640x640):	79.32 GFLOPs
Run Name:	yolov8m_testing_20251127_212555
Timestamp:	2025-11-27 21:38:09
Dataset:	bdd100k_yolo_limited - test split
Images Processed:	20000
Total Execution Time	113.08s
IoU Threshold:	0.5

Inference Performance

Metric	Value
Average Inference Time	2.38 ms per image
FPS (Frames Per Second)	420.94

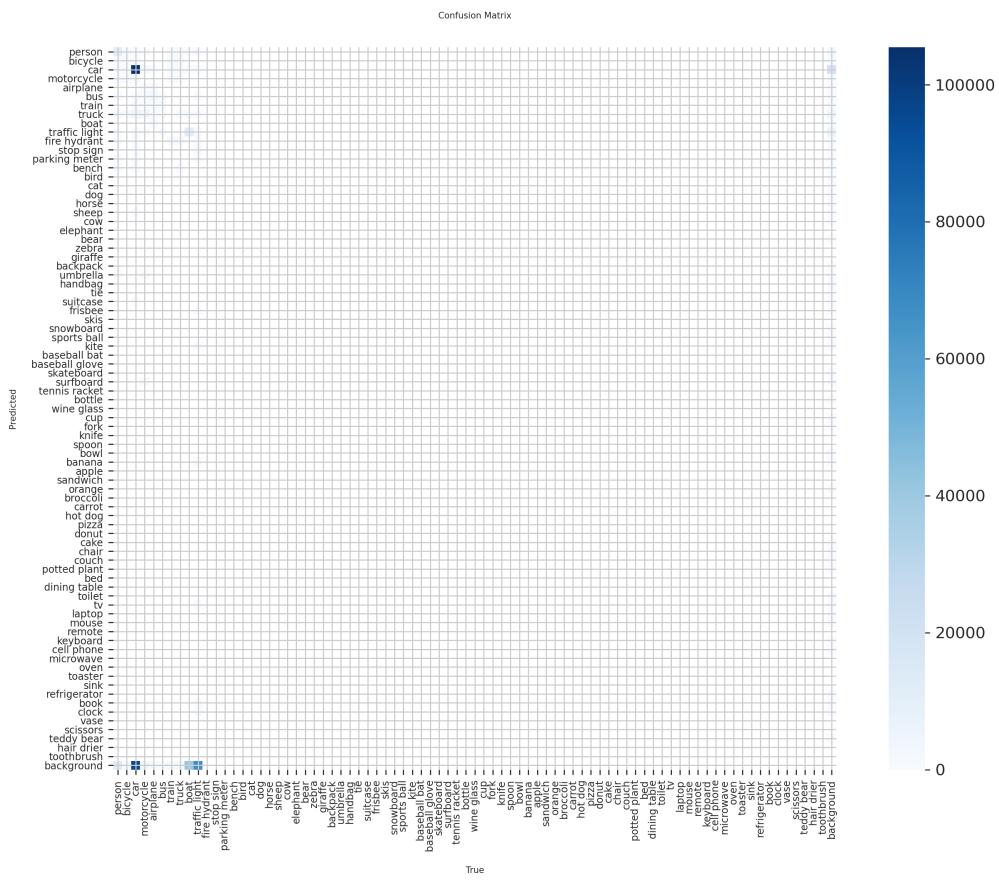
Note: Inference time includes preprocessing, model inference, and postprocessing.

Overall Accuracy Metrics

Metric	Value
Precision	0.3196
Recall	0.6765
F1-Score	0.4341
mAP@0.5	0.1225
mAP@0.5:0.95	0.0667
True Positives	117519
False Positives	250208
False Negatives	56205

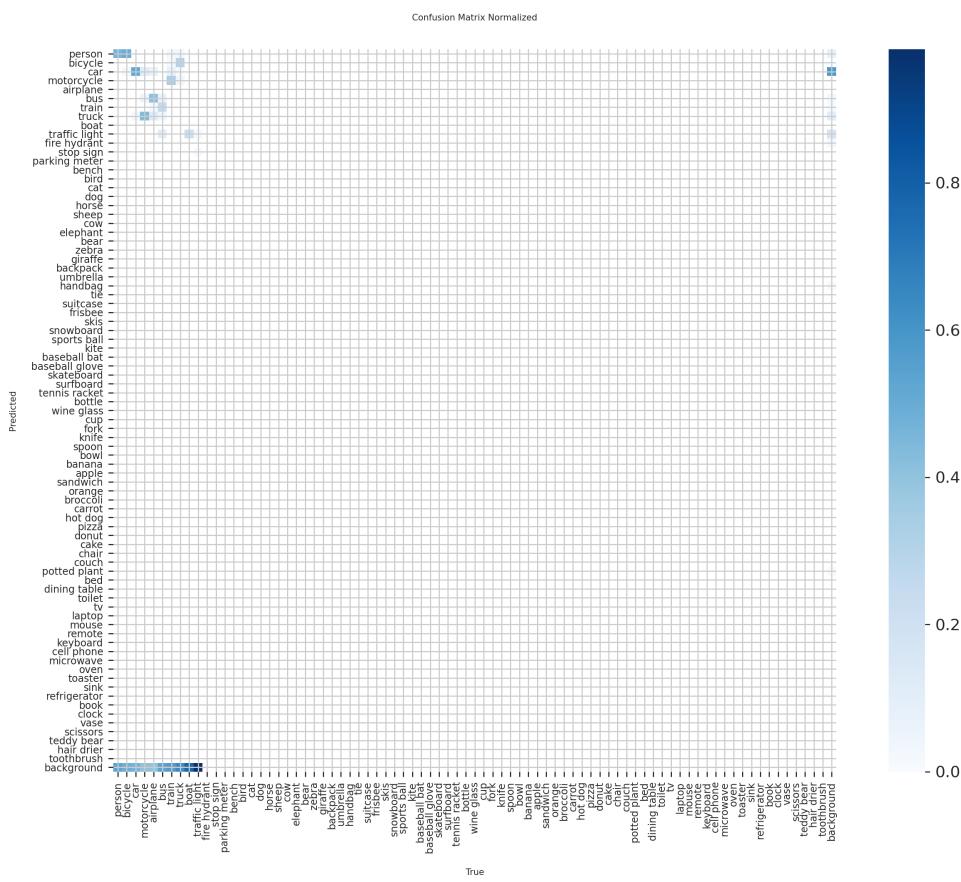
Confusion Matrices

Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

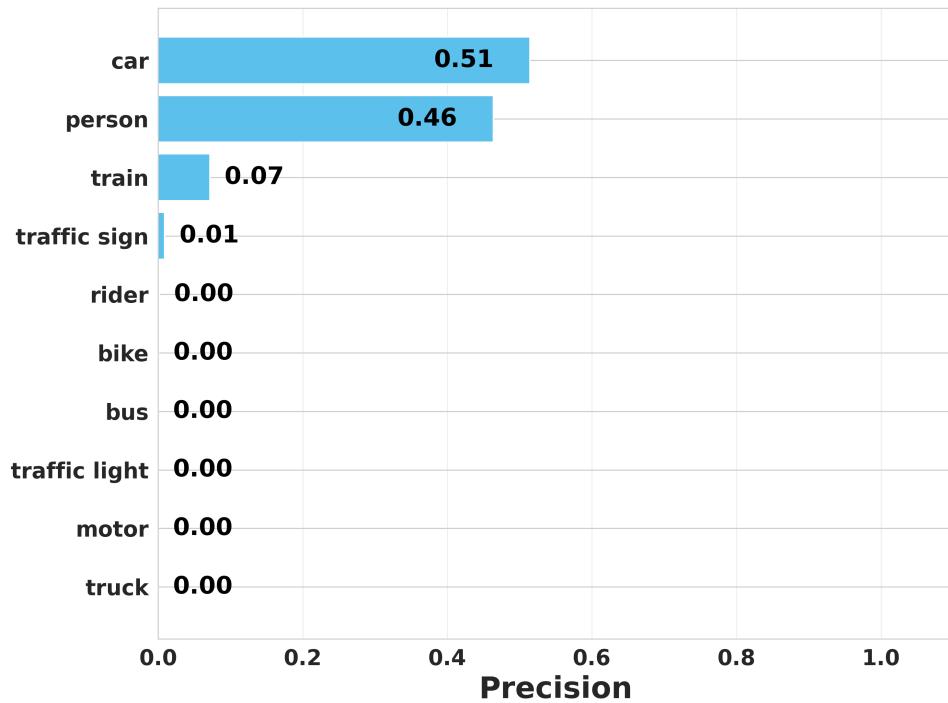
Confusion Matrix (Normalized)



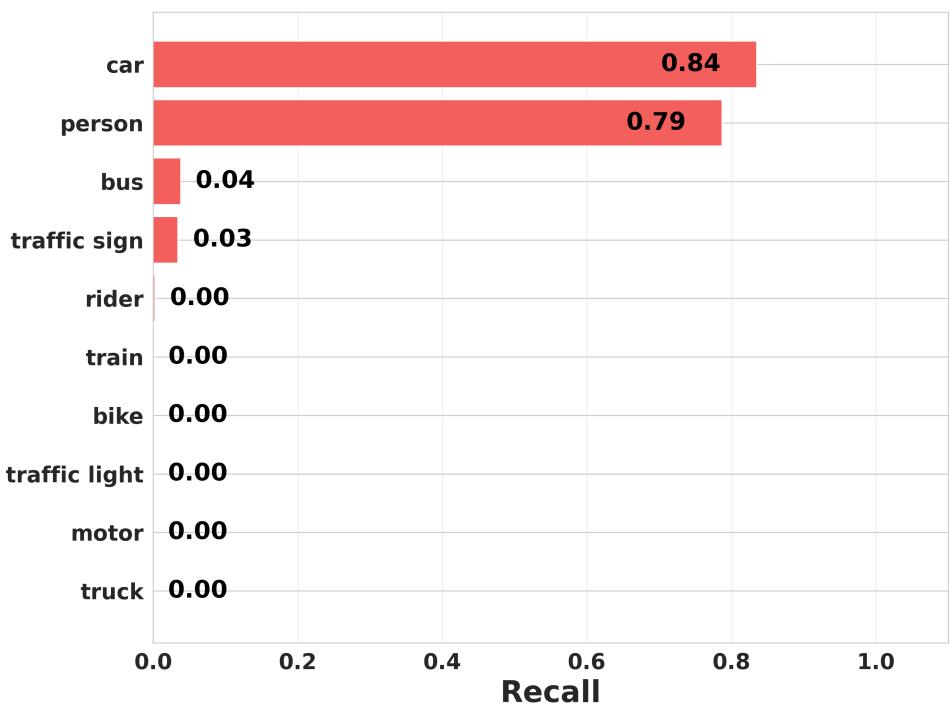
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

Performance Visualizations

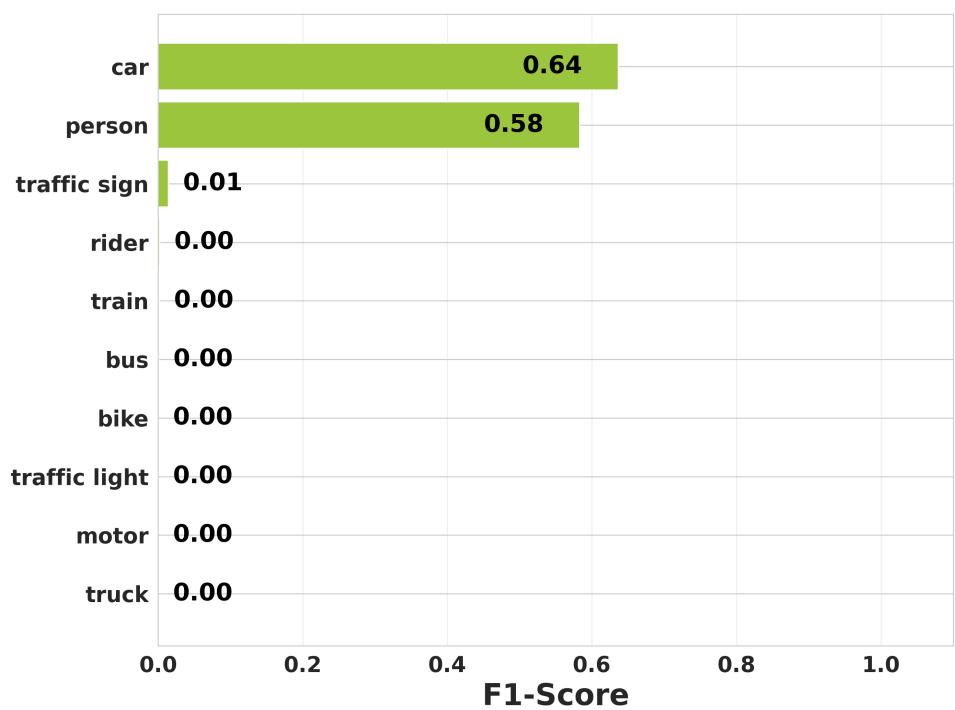
Precision by Class



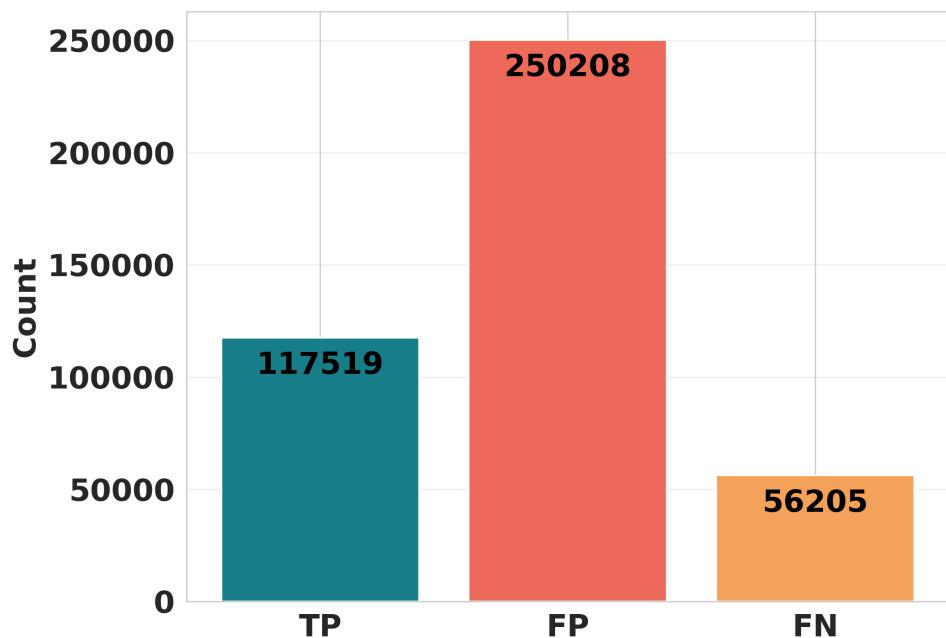
Recall by Class



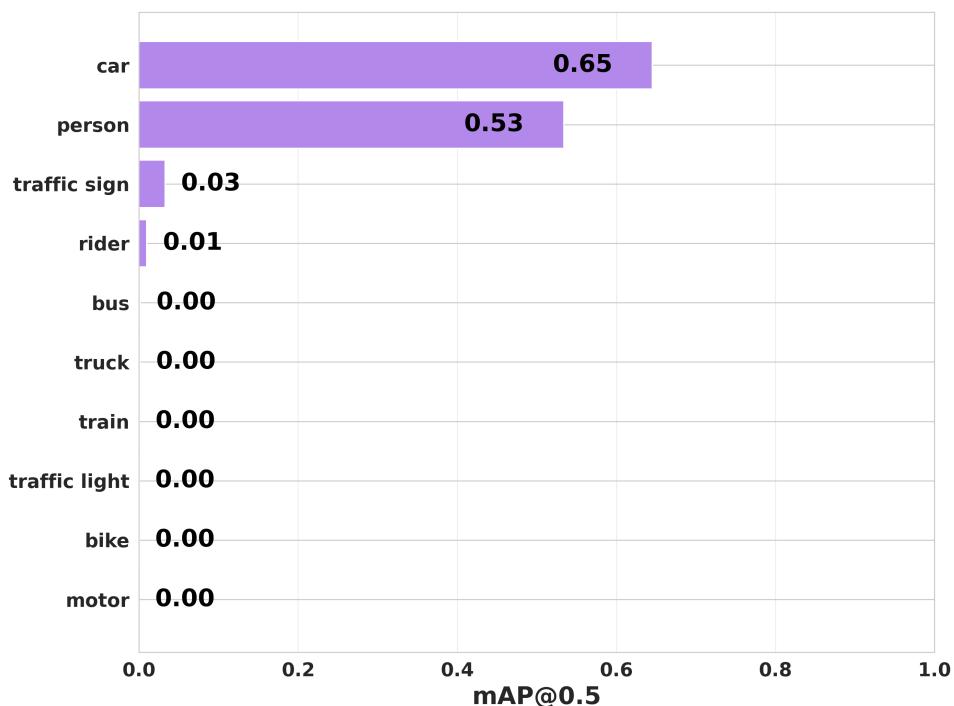
F1-Score by Class



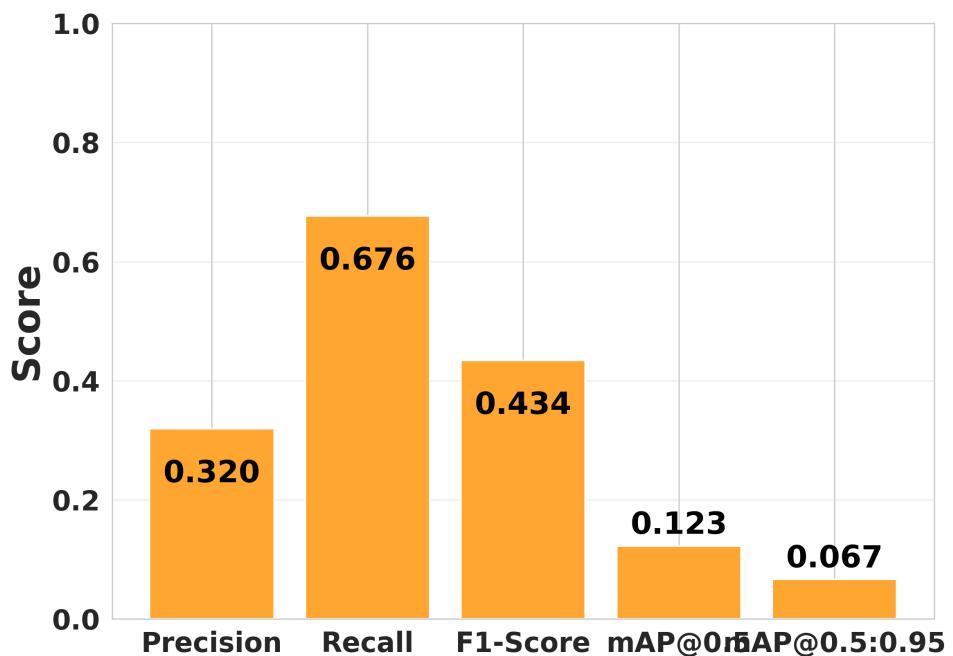
Overall Detection Outcomes



mAP@0.5 by Class



Overall Metrics



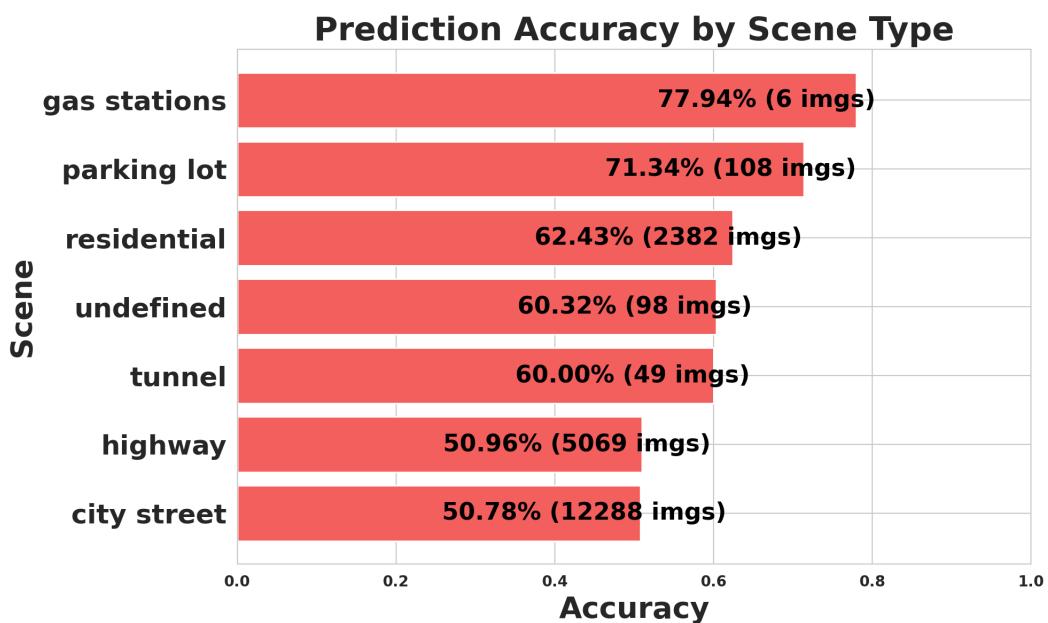
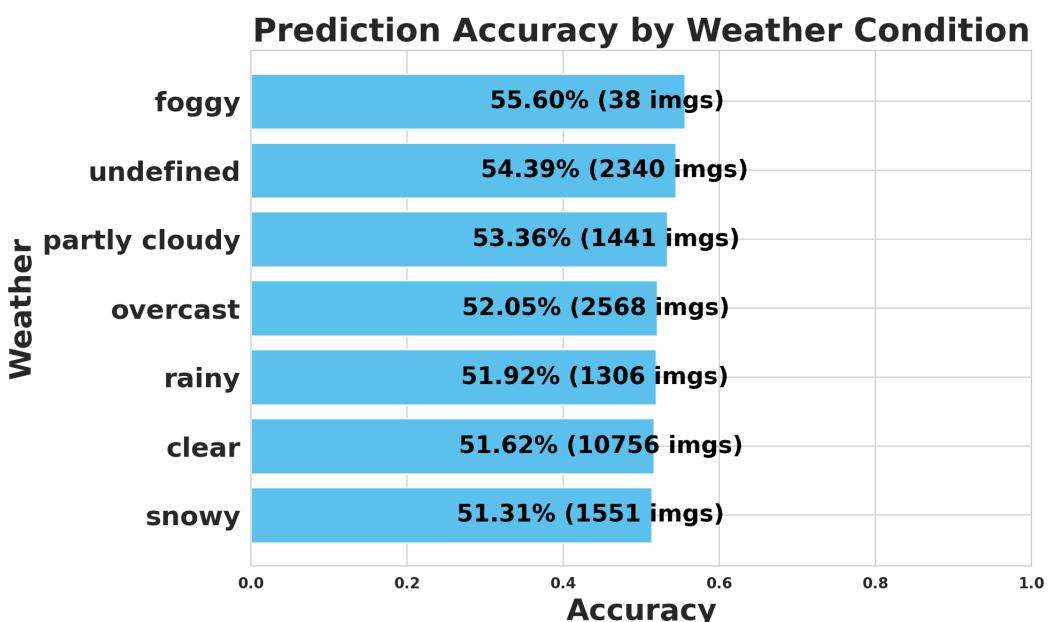
Per-Class Performance

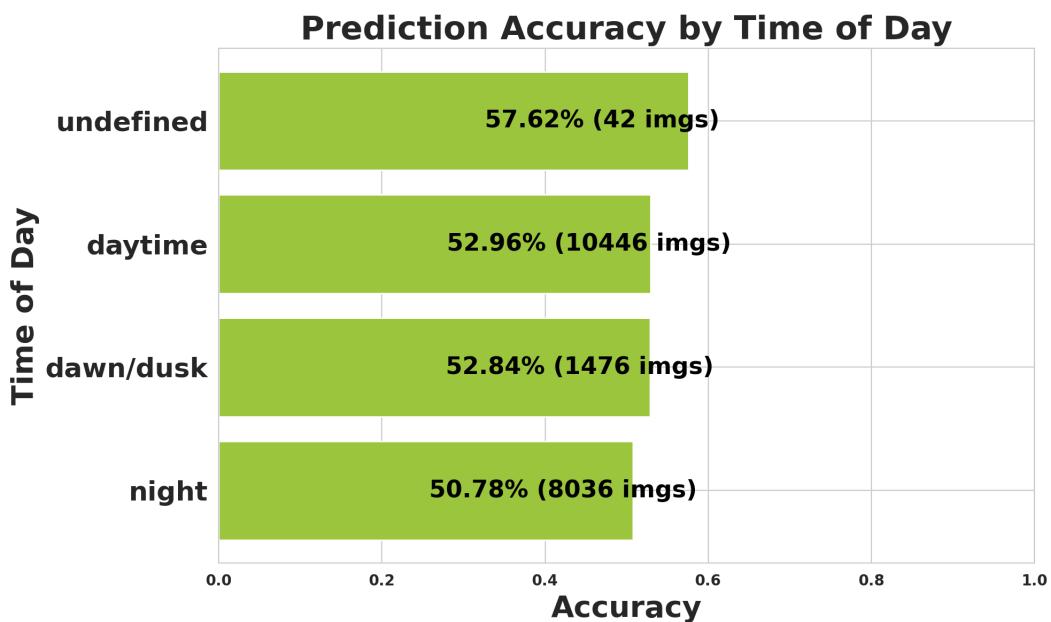
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	11425	13225	3087	0.4635	0.7873	0.5835	0.5339
rider	2	1292	734	0.0015	0.0027	0.0020	0.0096
car	105472	99677	20795	0.5141	0.8353	0.6365	0.6454
truck	0	8704	469	0.0000	0.0000	0.0000	0.0013
bus	1	3216	25	0.0003	0.0385	0.0006	0.0017
train	2	26	2883	0.0714	0.0007	0.0014	0.0003
motor	0	841	213	0.0000	0.0000	0.0000	0.0000
bike	1	1997	10752	0.0005	0.0001	0.0002	0.0001
traffic light	0	52840	13	0.0000	0.0000	0.0000	0.0002
traffic sign	616	68390	17234	0.0089	0.0345	0.0142	0.0327

Comprehensive Failure Analysis

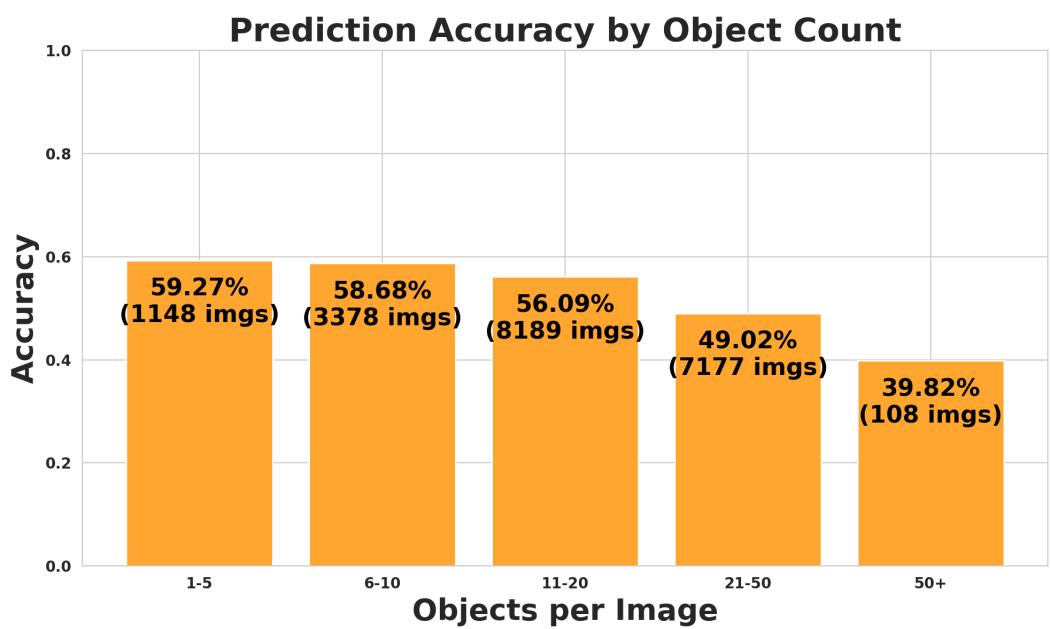
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	191856
Overall Accuracy	52.17%

Accuracy by Environmental Attributes

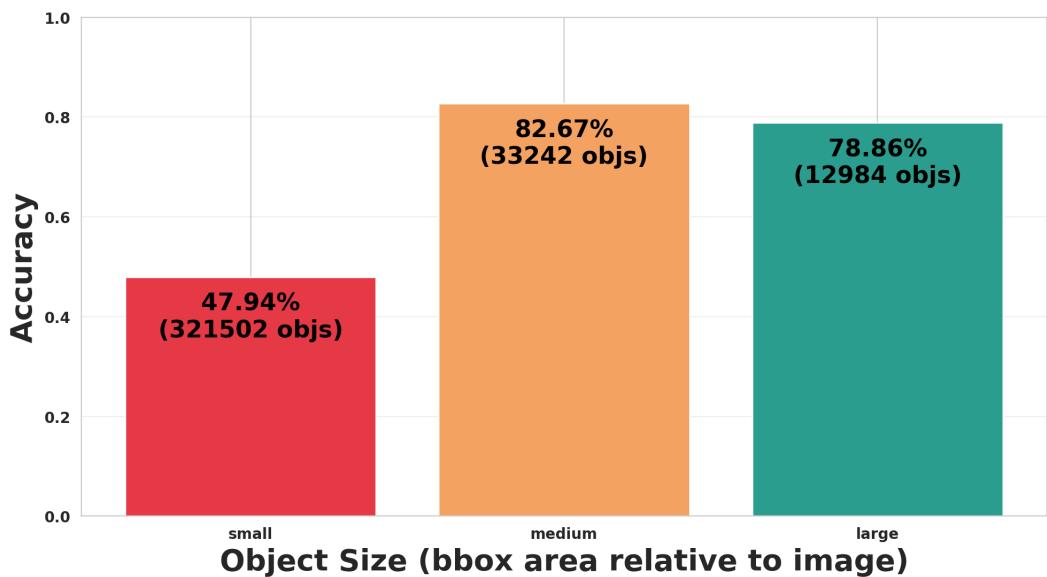




Accuracy by Object Characteristics

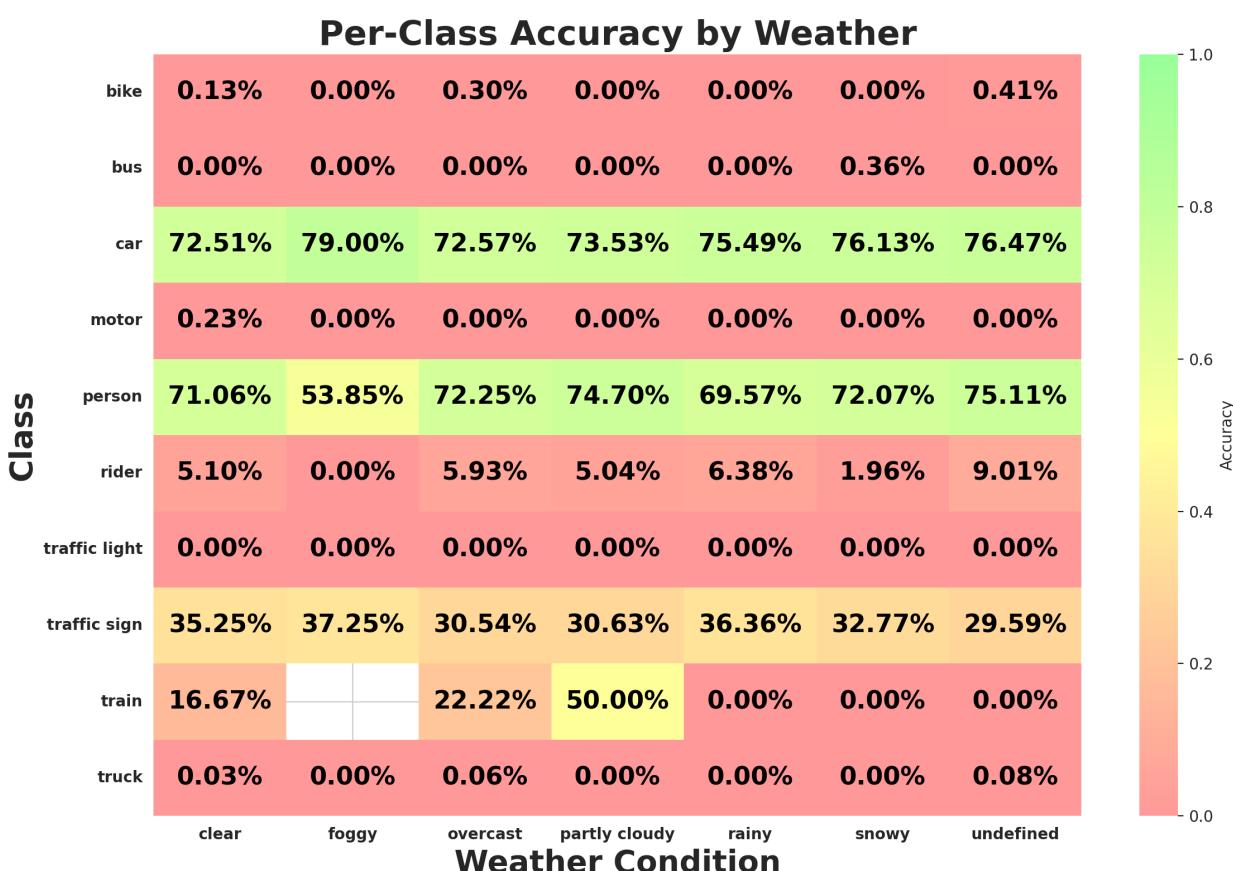
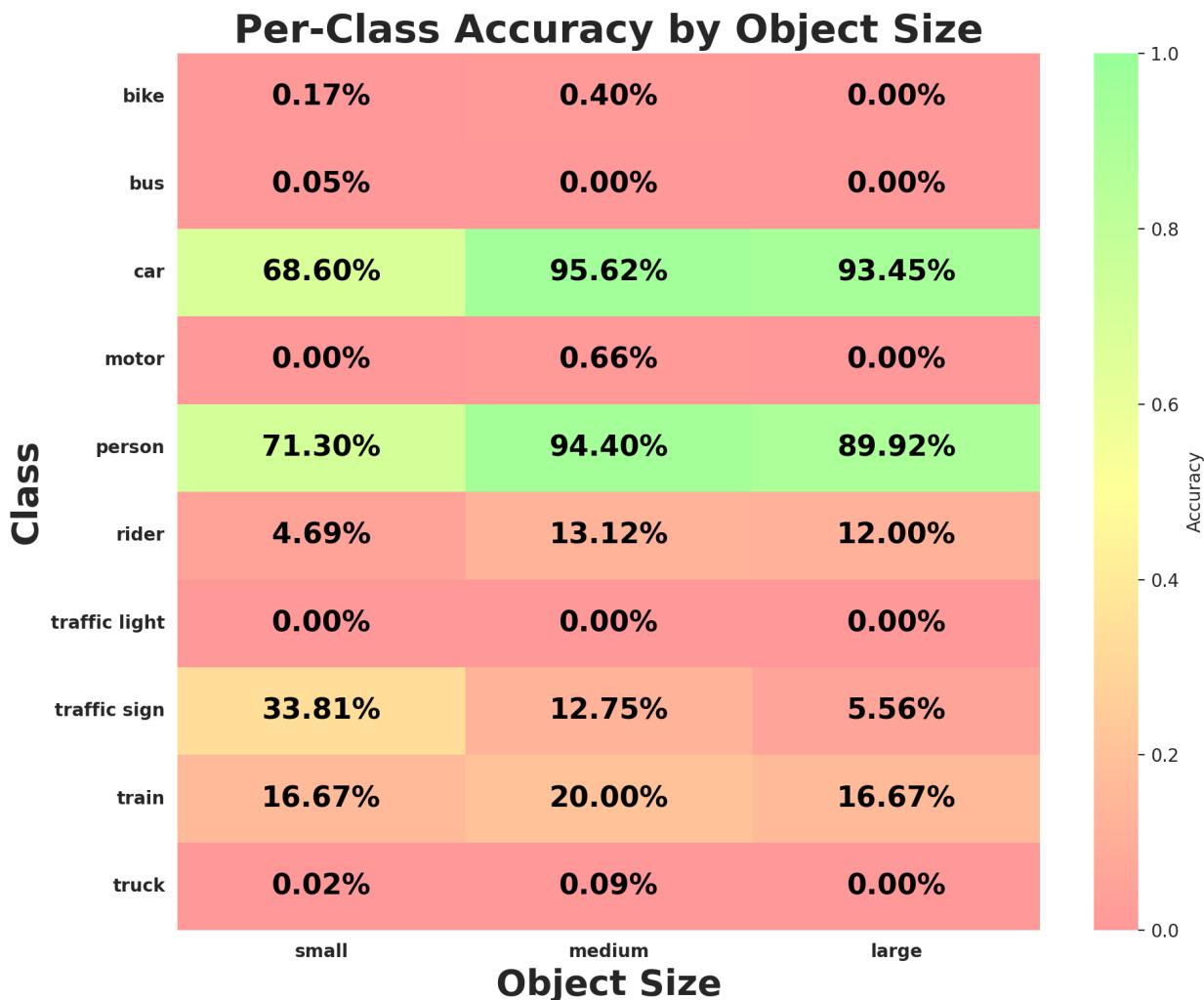


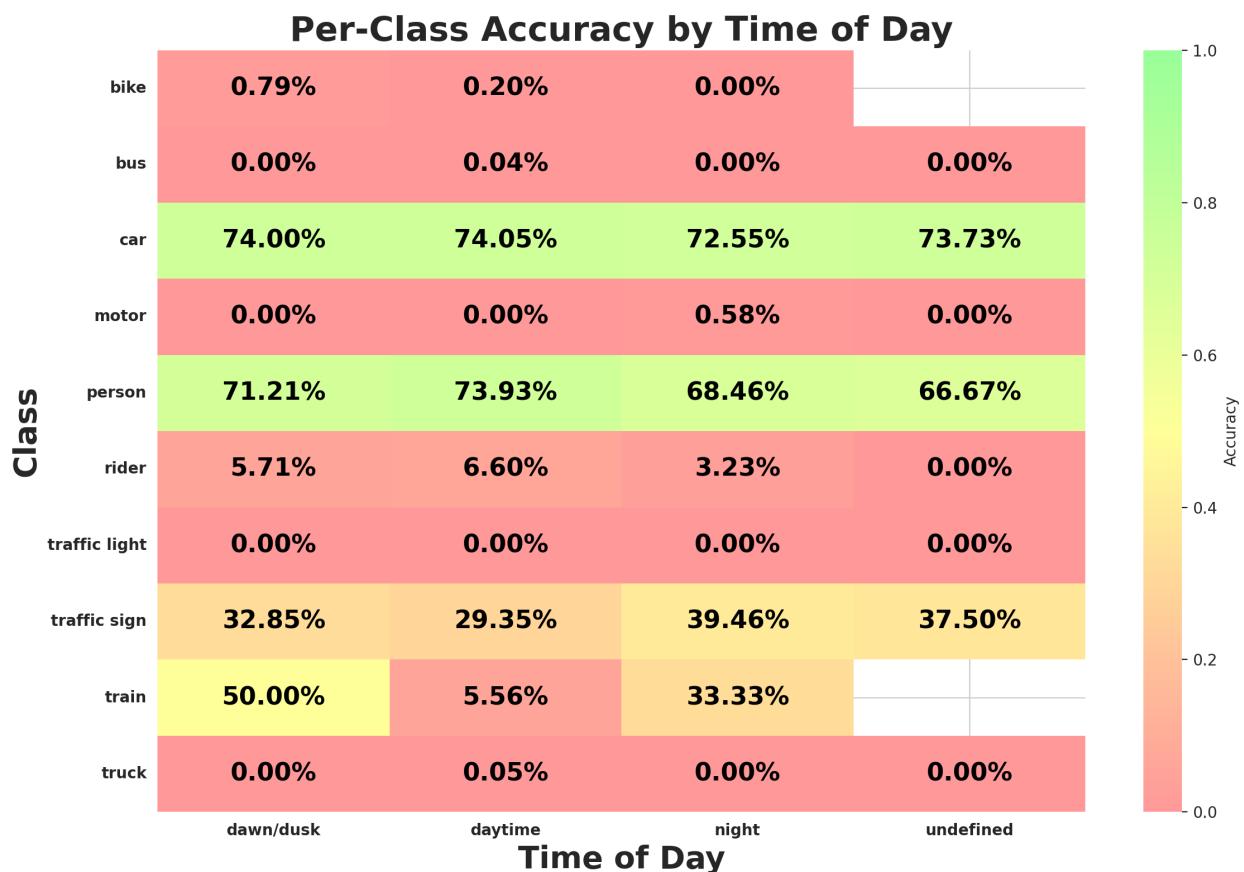
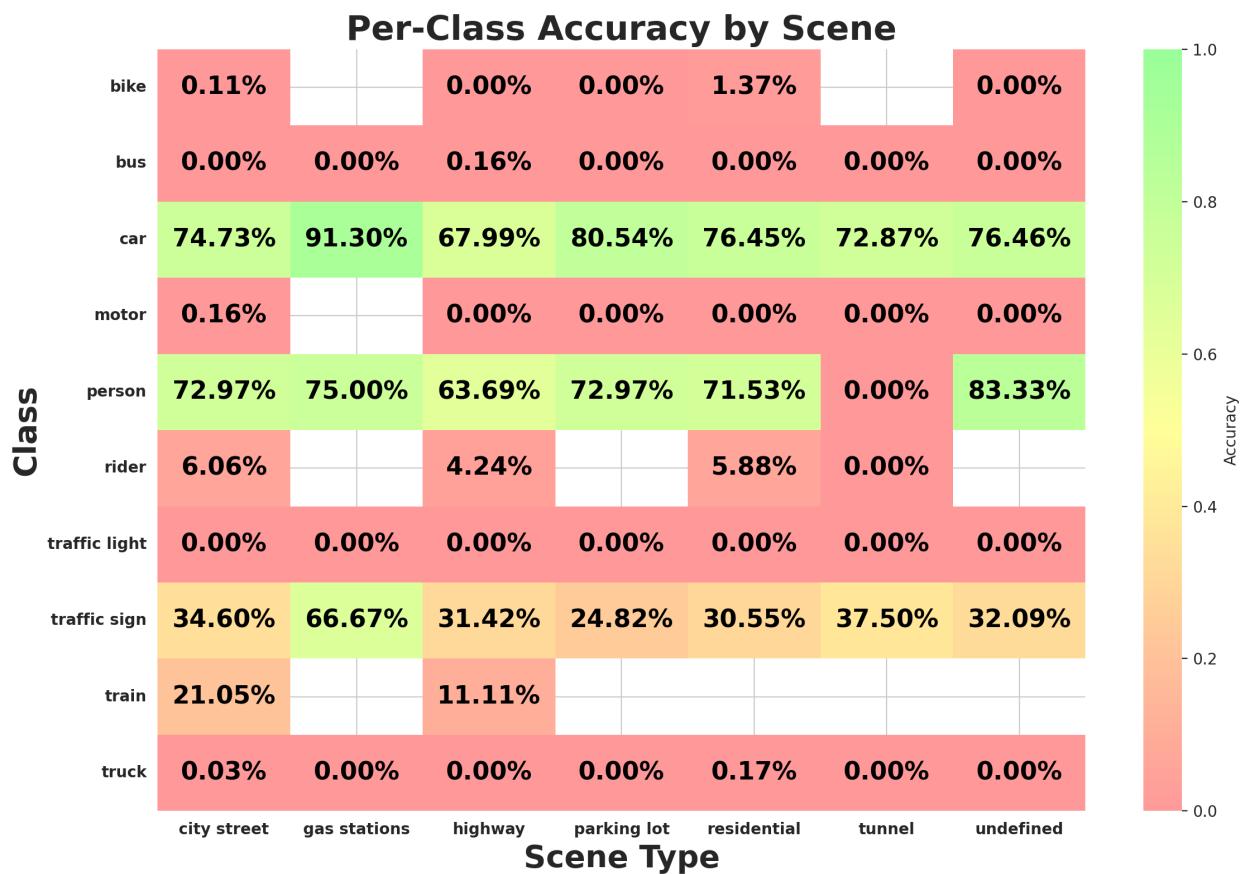
Prediction Accuracy by Object Size



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

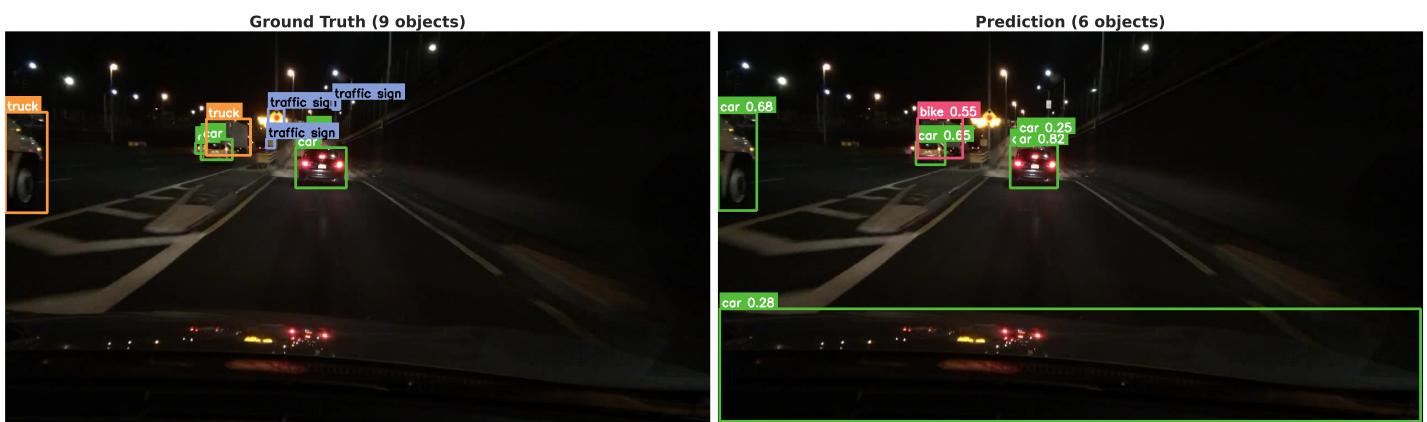
Sample #1: ea486e87-64934699.jpg

Ground Truth Objects: 5 | Predicted Objects: 2
Weather: clear | Scene: highway | Time of Day: night



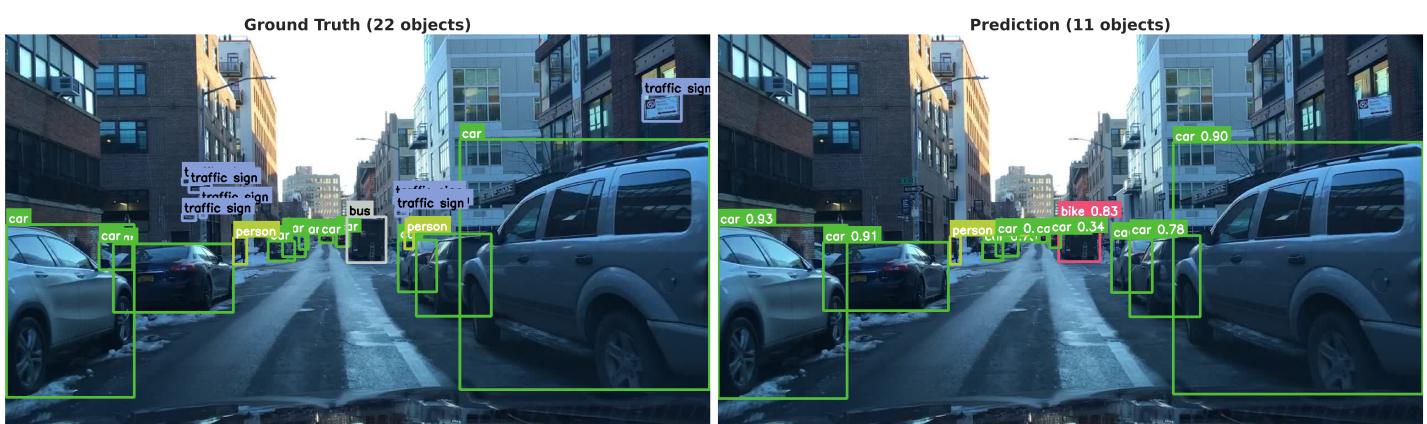
Sample #2: ef2a4126-9add4d07.jpg

Ground Truth Objects: 9 | Predicted Objects: 6
Weather: clear | Scene: highway | Time of Day: night



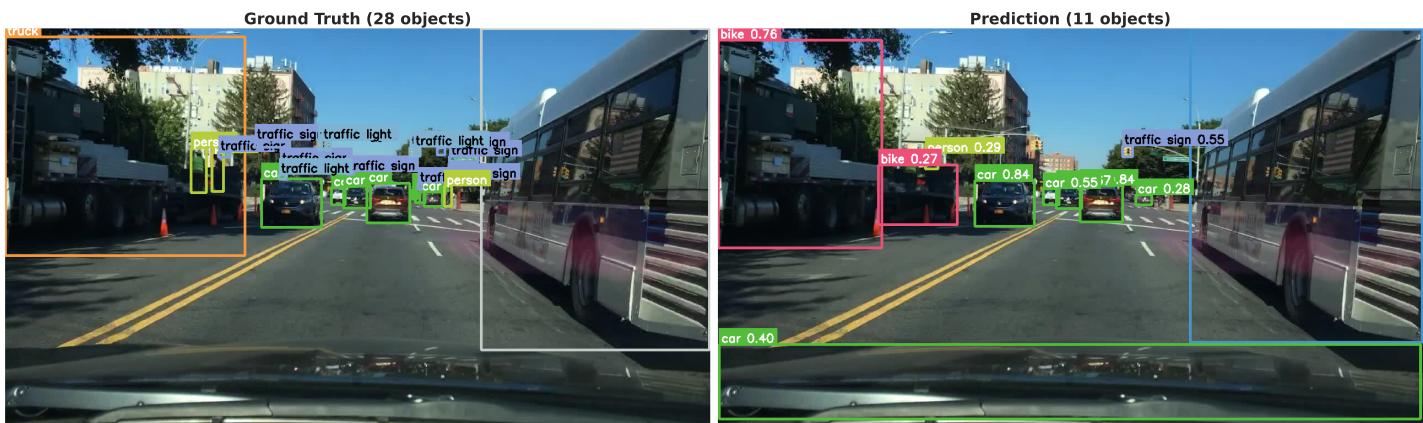
Sample #3: fa6f0cbe-78b6decd.jpg

Ground Truth Objects: 22 | Predicted Objects: 11
Weather: snowy | Scene: city street | Time of Day: dawn/dusk



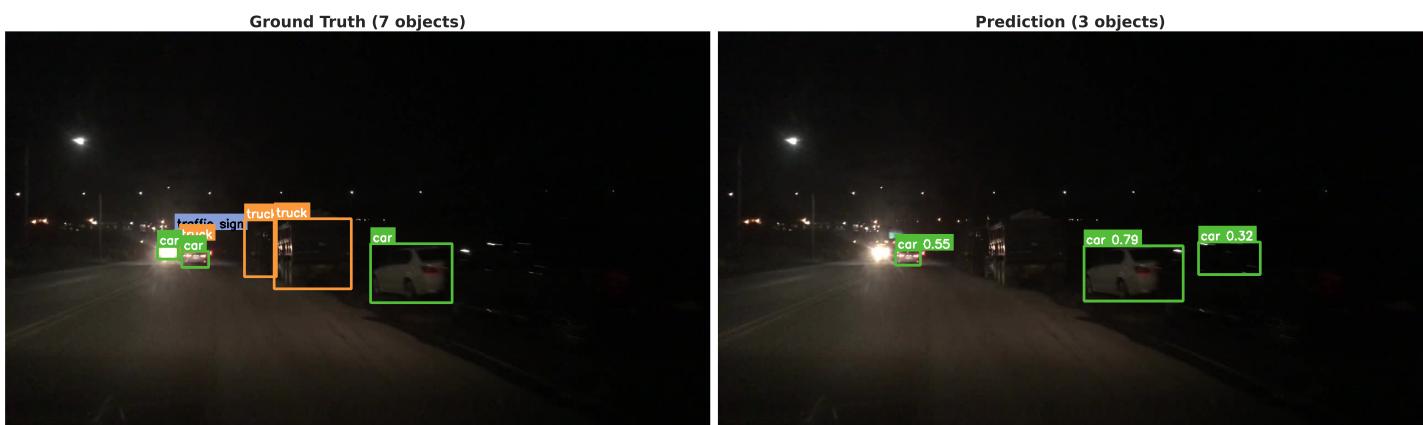
Sample #4: d23f5d6e-f041ed4e.jpg

Ground Truth Objects: 28 | Predicted Objects: 11
Weather: clear | Scene: city street | Time of Day: daytime



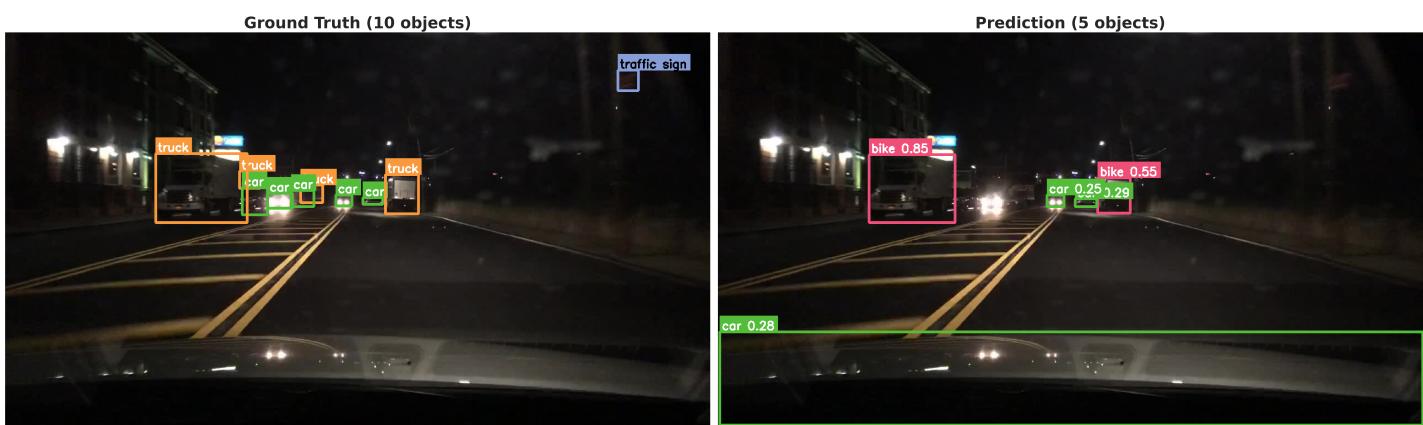
Sample #5: ed3c1812-962419ef.jpg

Ground Truth Objects: 7 | Predicted Objects: 3
Weather: clear | Scene: city street | Time of Day: night



Sample #6: dfe4d2cd-2fa96a68.jpg

Ground Truth Objects: 10 | Predicted Objects: 5
Weather: clear | Scene: city street | Time of Day: night



Generated by YOLO Testing Pipeline