

YOLO Model Testing Report

Model:	yolo12x
Model Size:	113.8 MB
Parameters:	59.2 M
FLOPs (640x640):	200.33 GFLOPs
Run Name:	yolo12x_testing_20251128_014537
Timestamp:	2025-11-28 02:03:49
Dataset:	bdd100k_yolo_limited - test split
Images Processed:	20000
Total Execution Time	188.20s
IoU Threshold:	0.5

Inference Performance

Metric	Value
Average Inference Time	6.03 ms per image
FPS (Frames Per Second)	165.76

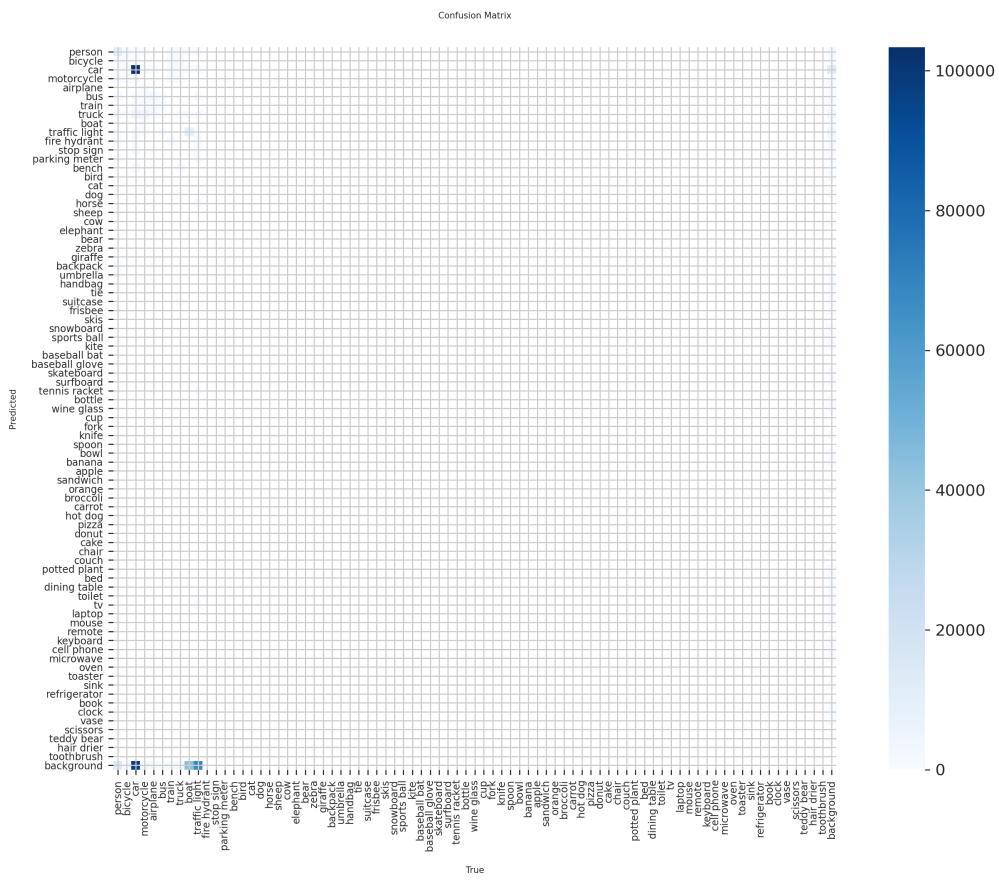
Note: Inference time includes preprocessing, model inference, and postprocessing.

Overall Accuracy Metrics

Metric	Value
Precision	0.3104
Recall	0.7300
F1-Score	0.4356
mAP@0.5	0.1306
mAP@0.5:0.95	0.0733
True Positives	114139
False Positives	253588
False Negatives	42220

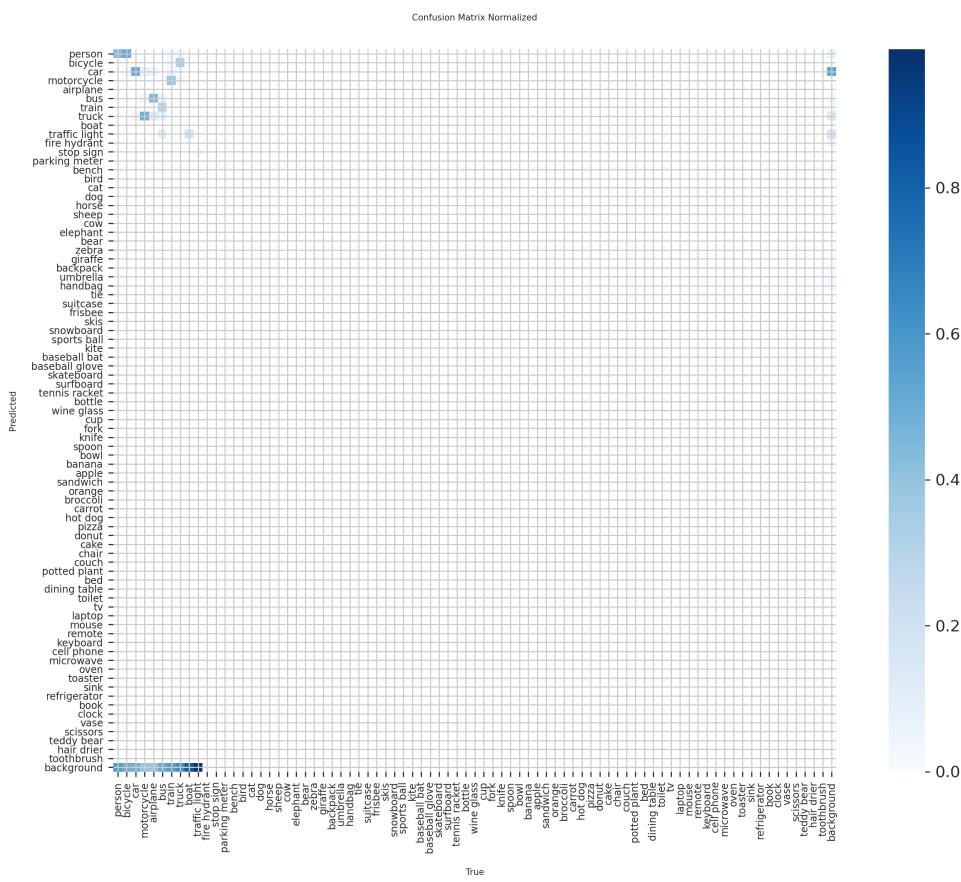
Confusion Matrices

Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

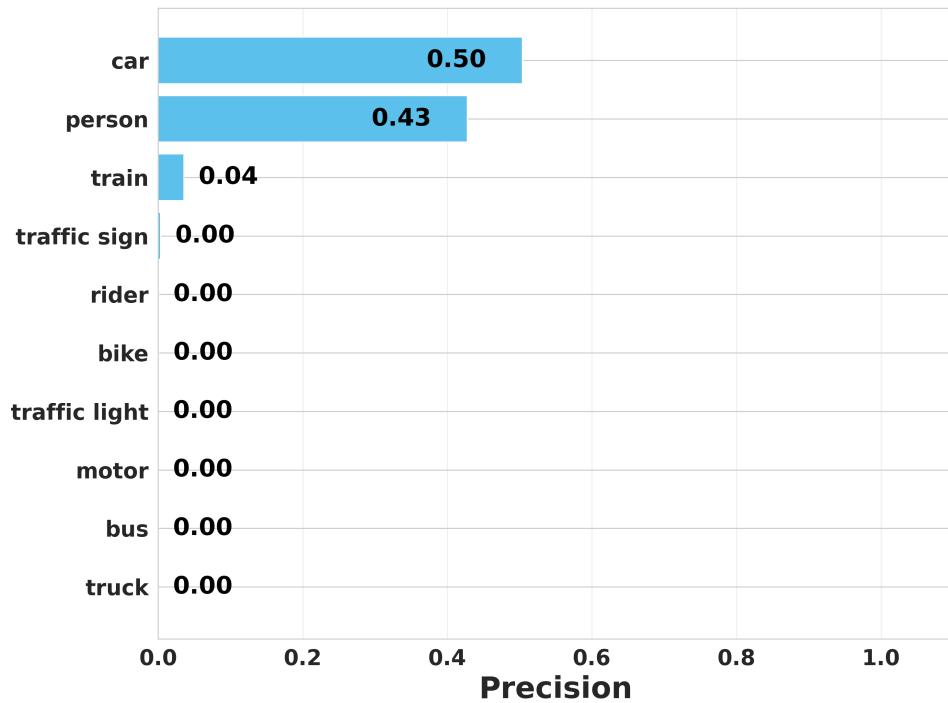
Confusion Matrix (Normalized)



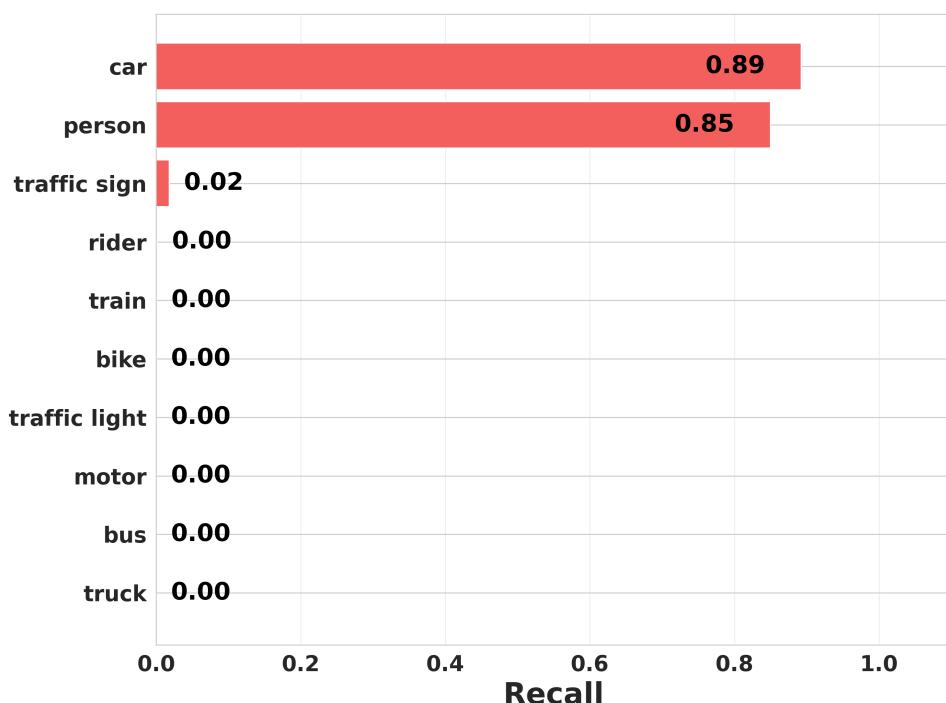
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

Performance Visualizations

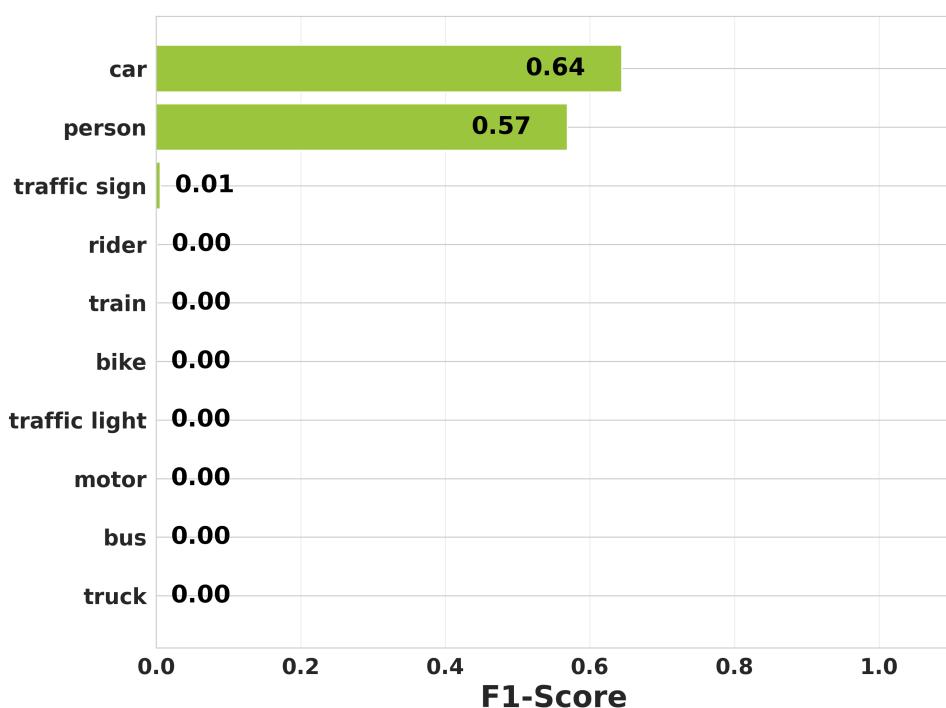
Precision by Class



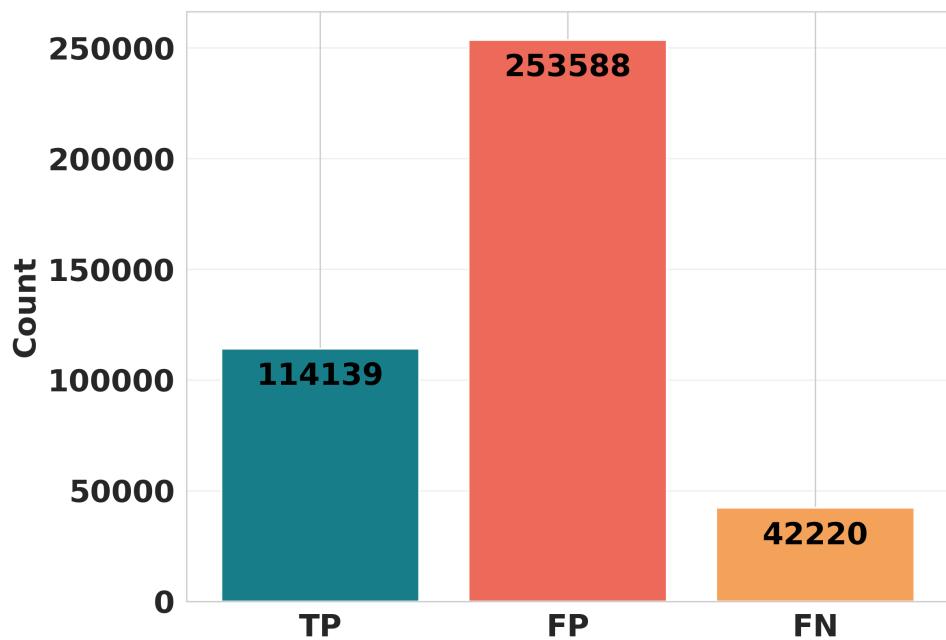
Recall by Class



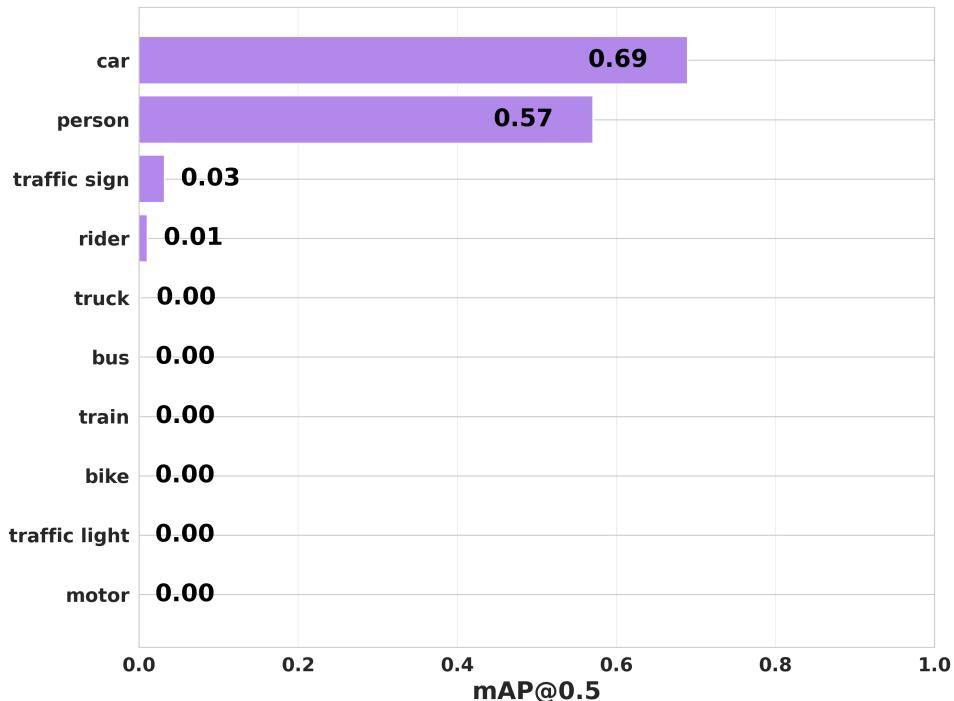
F1-Score by Class



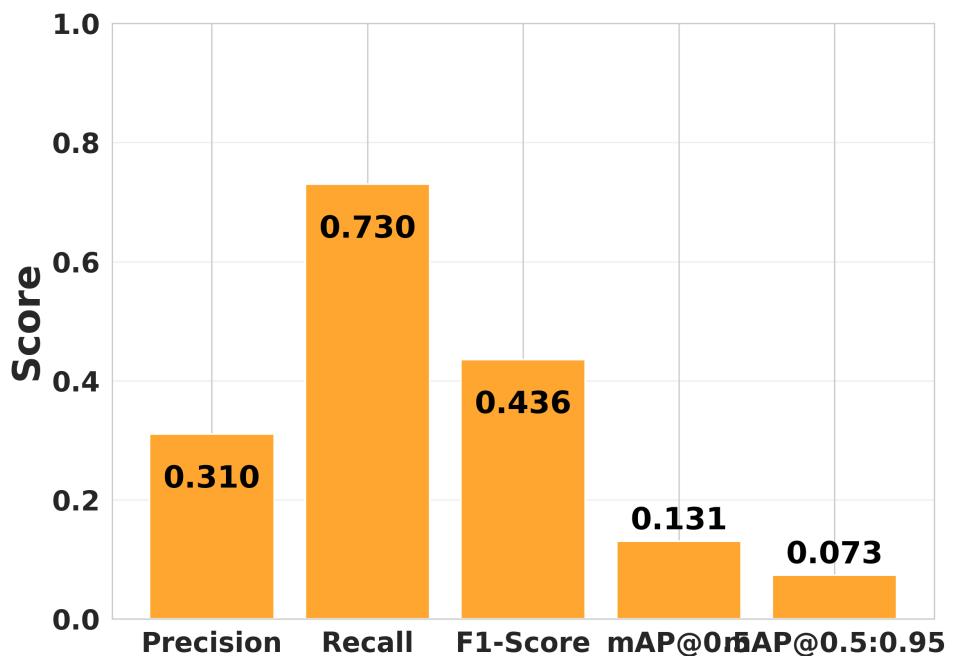
Overall Detection Outcomes



mAP@0.5 by Class



Overall Metrics



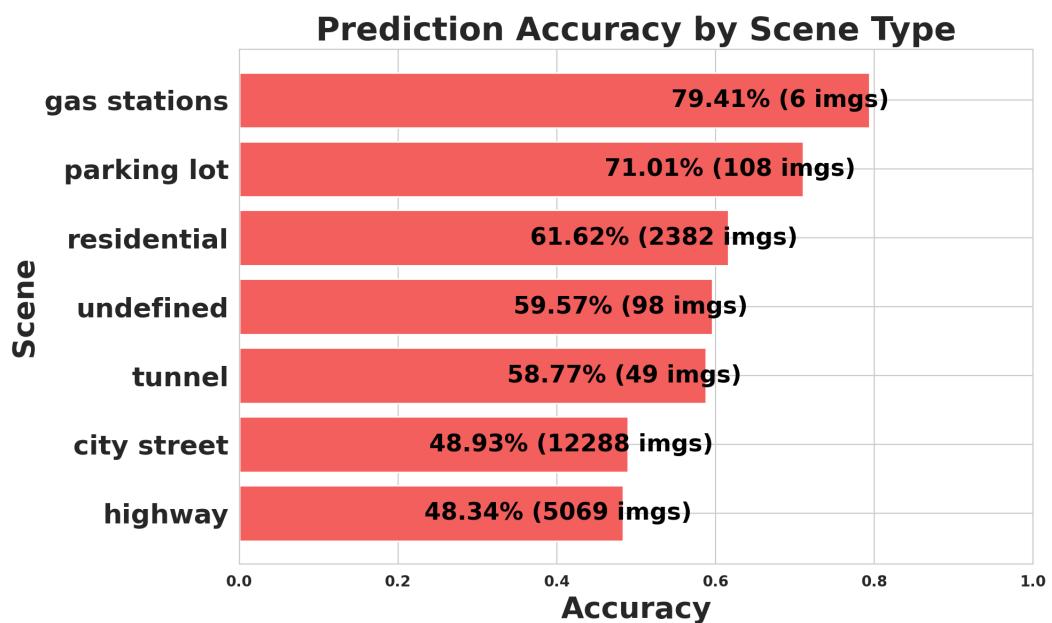
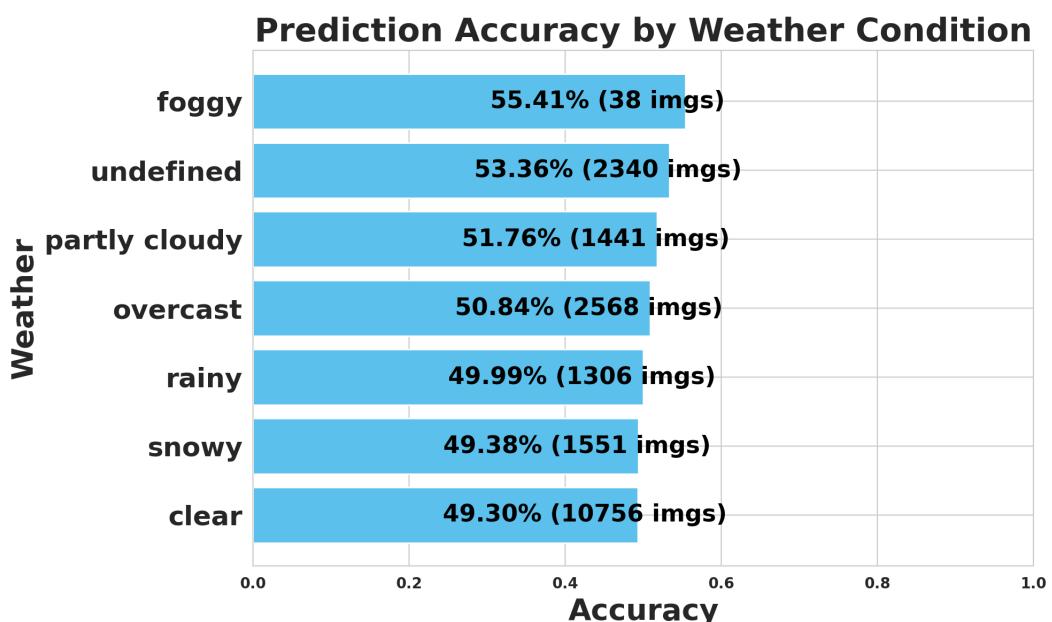
Per-Class Performance

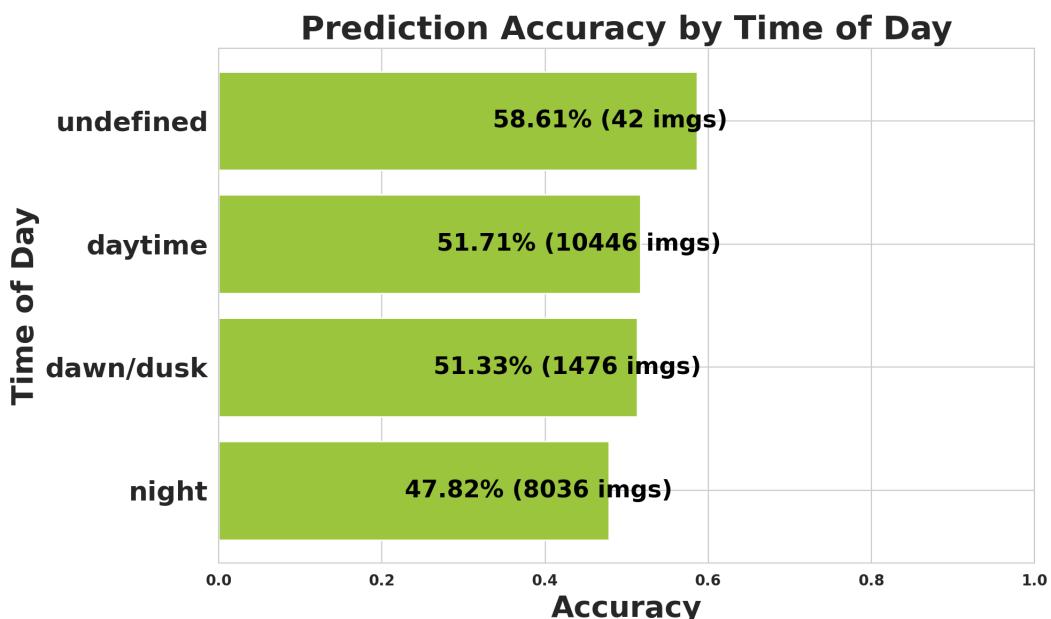
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	10544	14106	1859	0.4277	0.8501	0.5691	0.5707
rider	1	1293	753	0.0008	0.0013	0.0010	0.0107
car	103358	101791	12460	0.5038	0.8924	0.6440	0.6897
truck	0	8704	468	0.0000	0.0000	0.0000	0.0016
bus	0	3217	22	0.0000	0.0000	0.0000	0.0005
train	1	27	2641	0.0357	0.0004	0.0007	0.0005
motor	0	841	183	0.0000	0.0000	0.0000	0.0000
bike	1	1997	11055	0.0005	0.0001	0.0002	0.0001
traffic light	0	52840	10	0.0000	0.0000	0.0000	0.0000
traffic sign	234	68772	12769	0.0034	0.0180	0.0057	0.0320

Comprehensive Failure Analysis

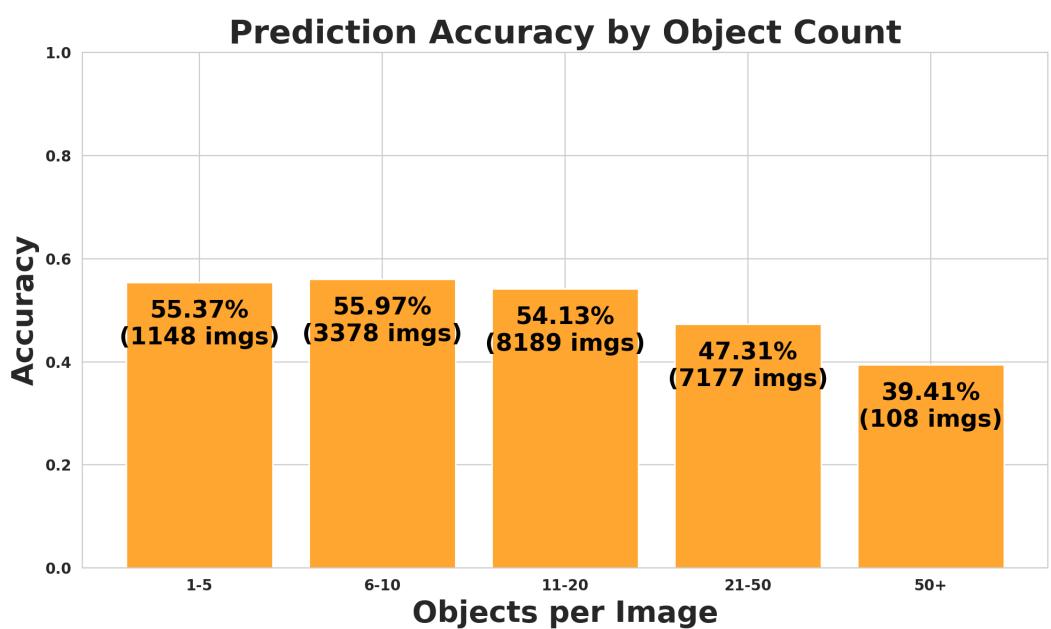
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	184945
Overall Accuracy	50.29%

Accuracy by Environmental Attributes

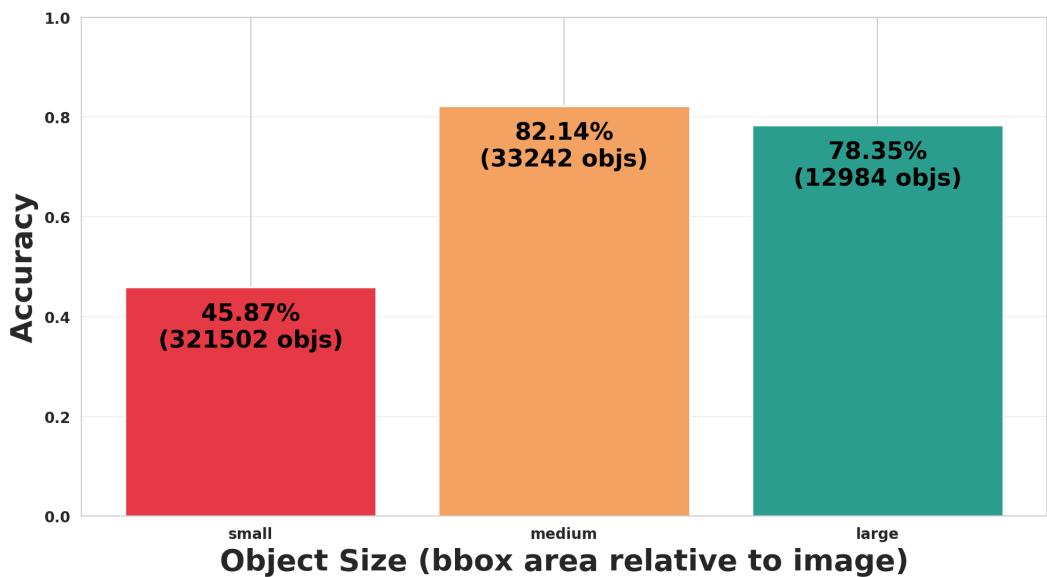




Accuracy by Object Characteristics

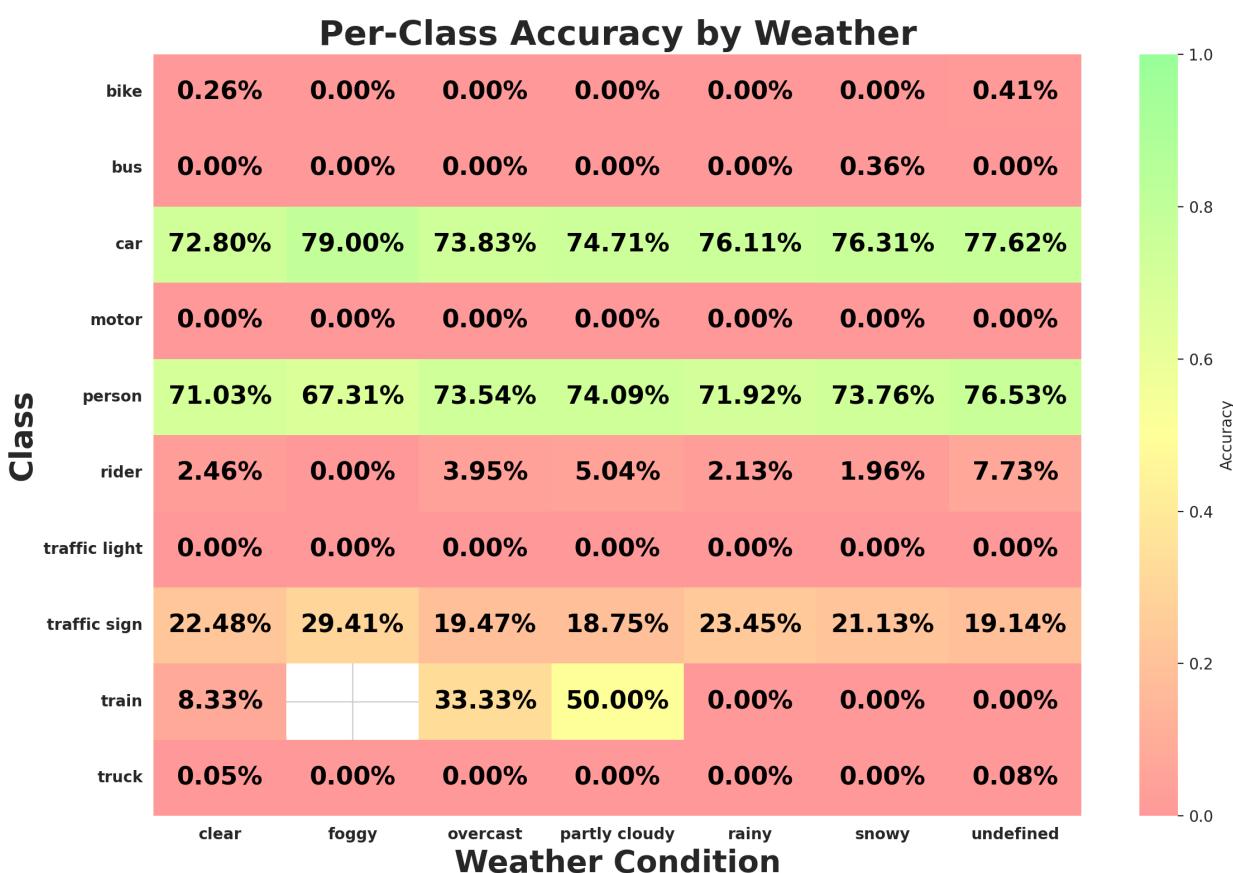
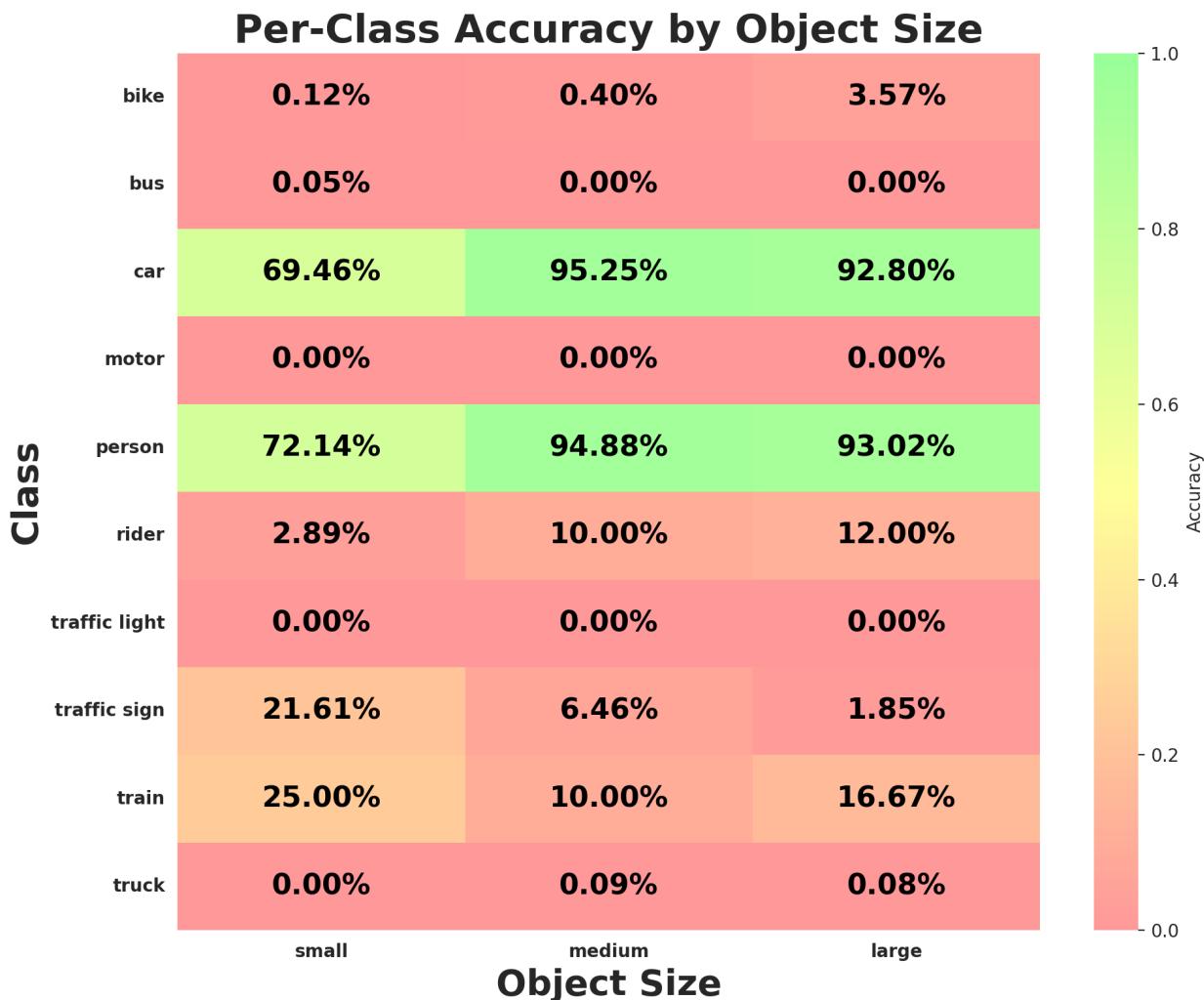


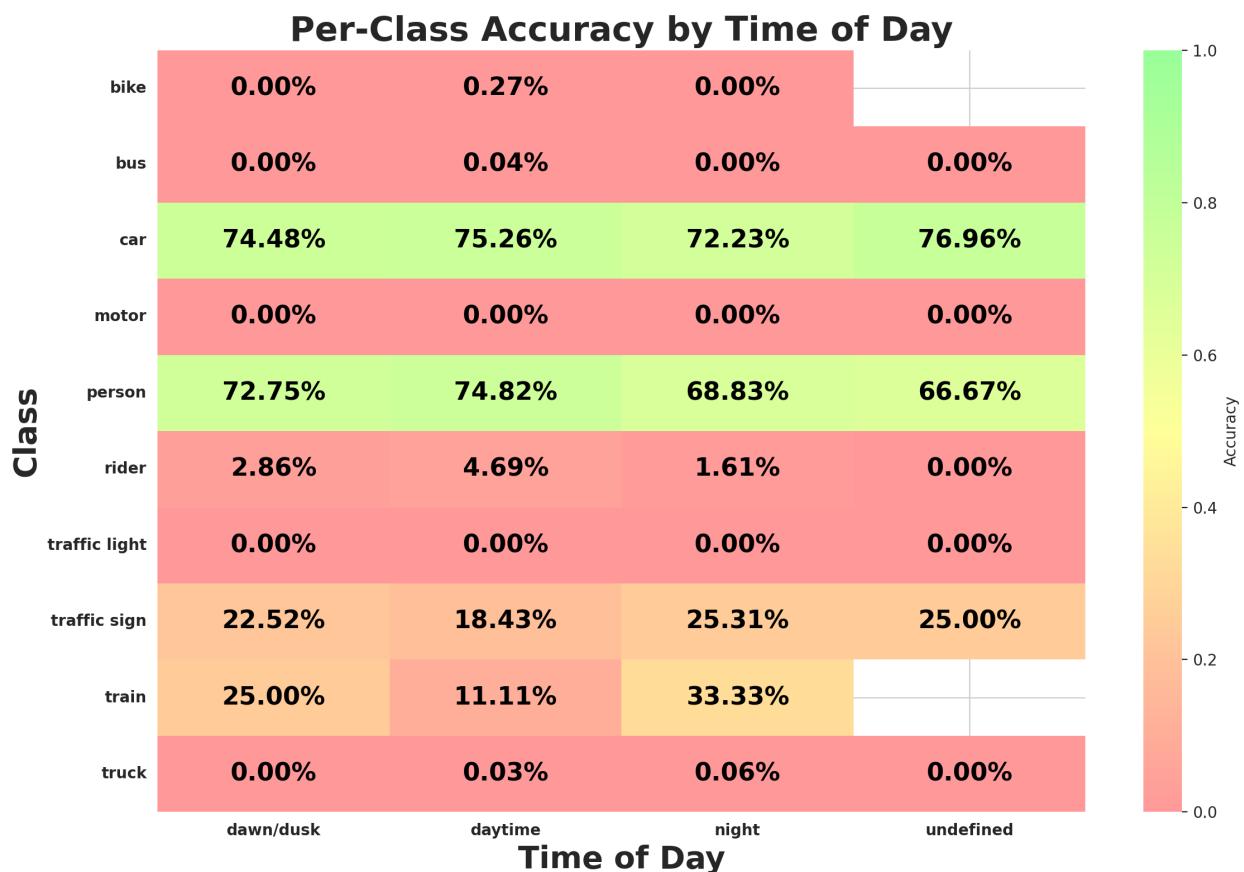
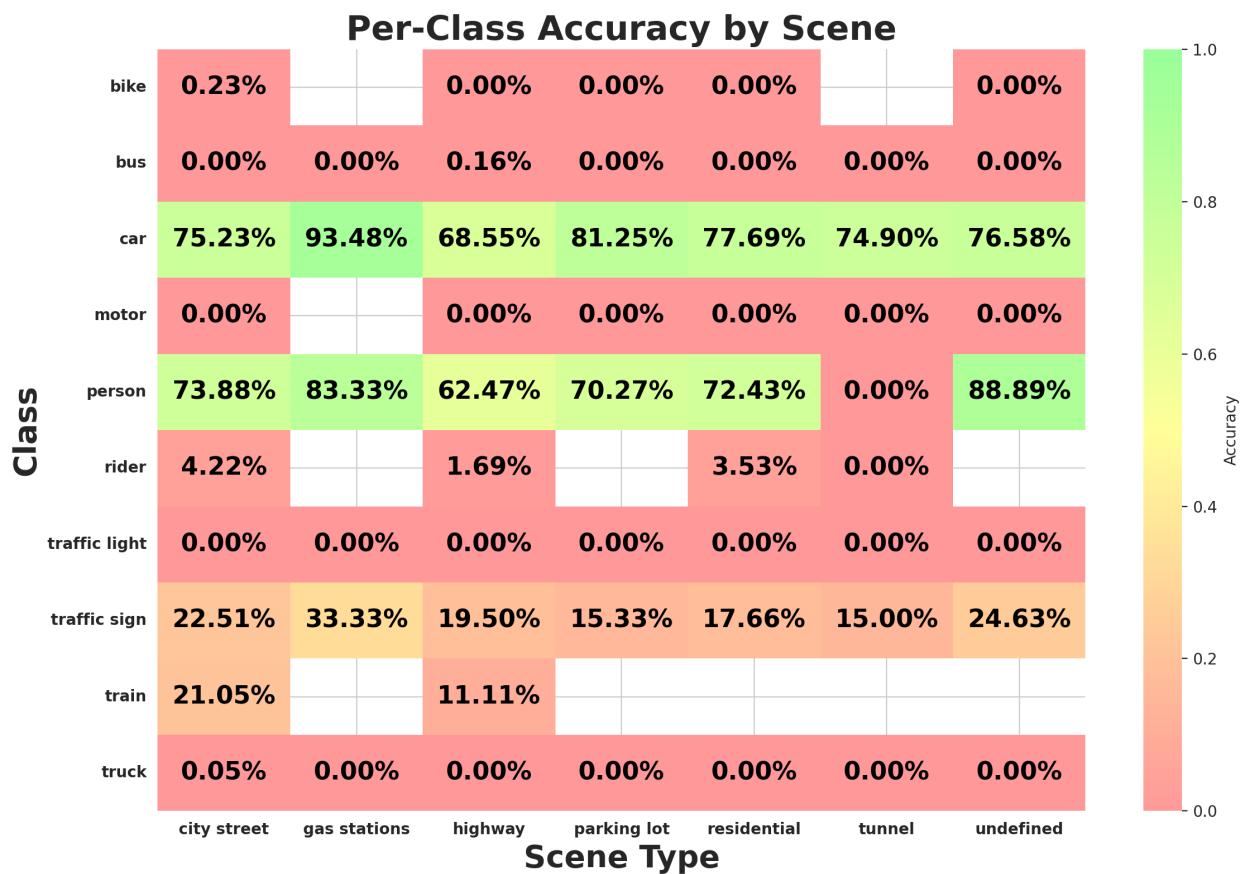
Prediction Accuracy by Object Size



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: d2c4ceac-03dc4f58.jpg

Ground Truth Objects: 10 | Predicted Objects: 5

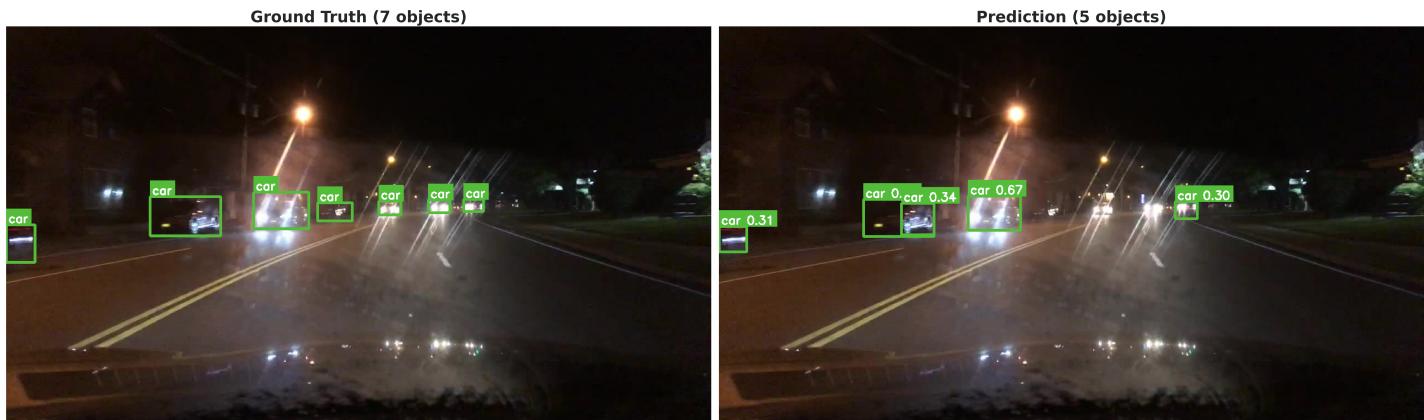
Weather: snowy | Scene: city street | Time of Day: dawn/dusk



Sample #2: f3afbbbeb-35db2058.jpg

Ground Truth Objects: 7 | Predicted Objects: 5

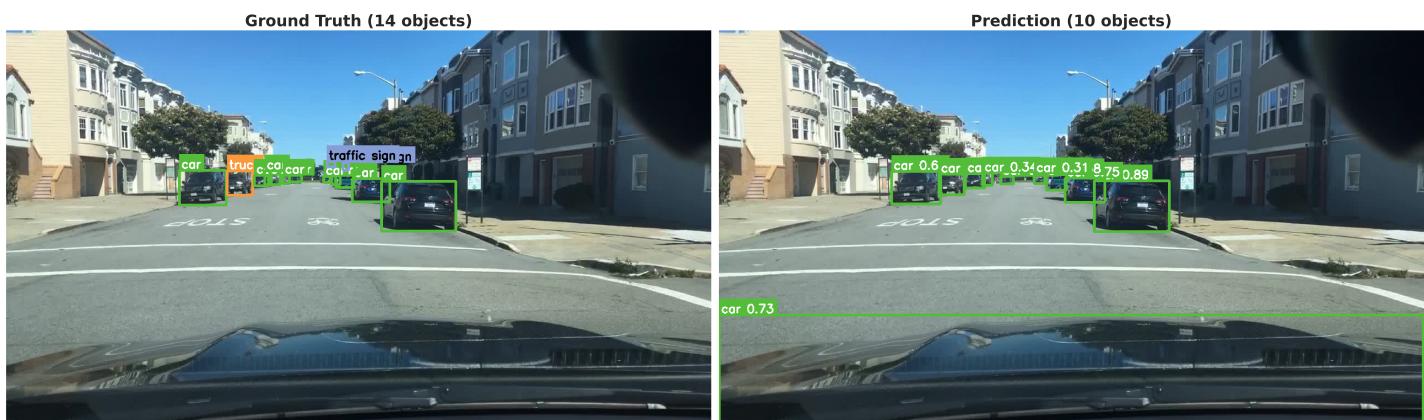
Weather: clear | Scene: city street | Time of Day: night



Sample #3: db24bc52-04a0d171.jpg

Ground Truth Objects: 14 | Predicted Objects: 10

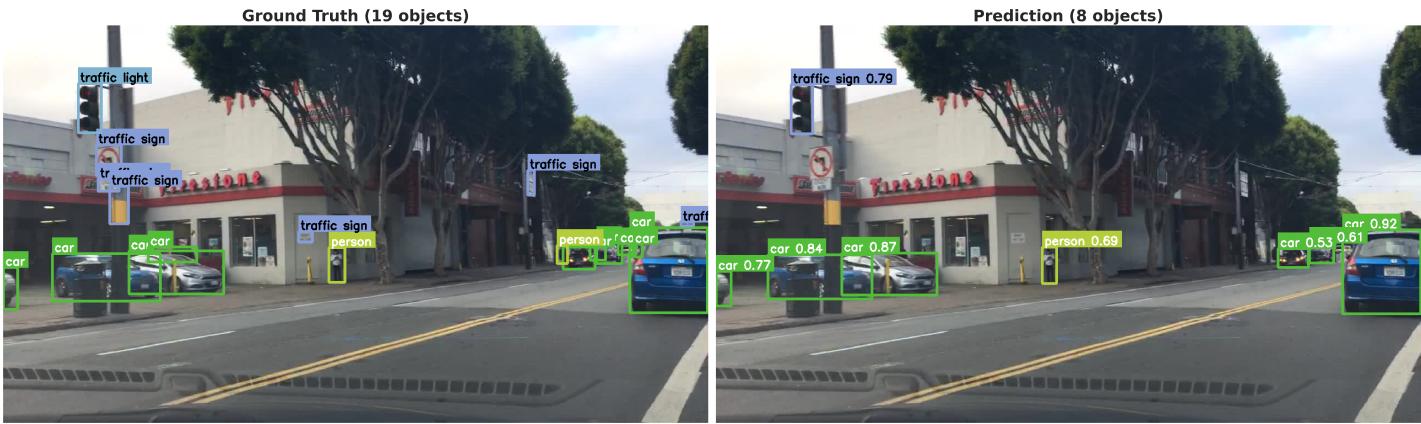
Weather: clear | Scene: residential | Time of Day: daytime



Sample #4: fa3a71c8-5202228f.jpg

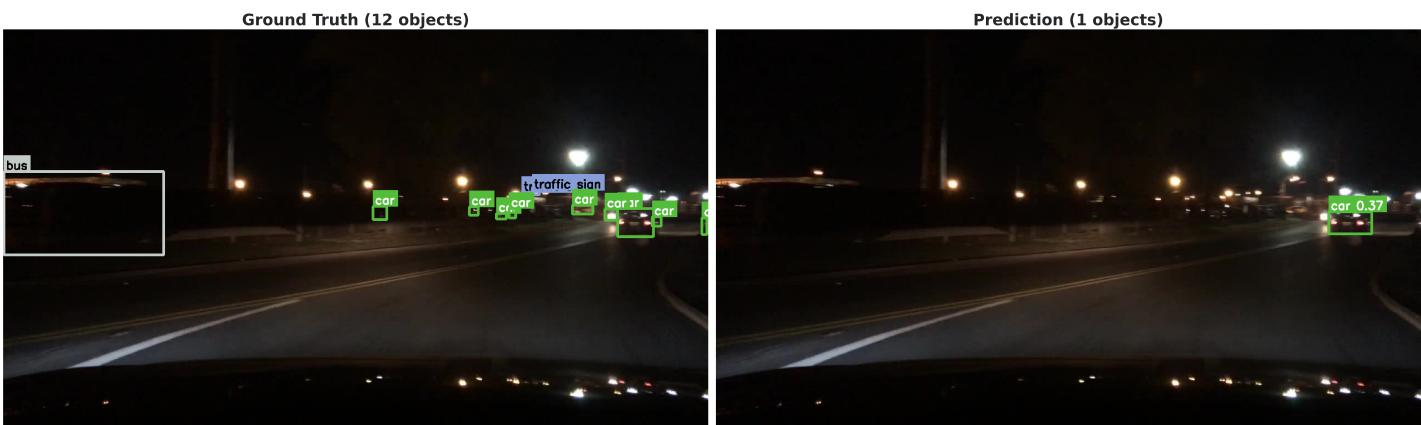
Ground Truth Objects: 19 | Predicted Objects: 8

Weather: partly cloudy | Scene: city street | Time of Day: daytime



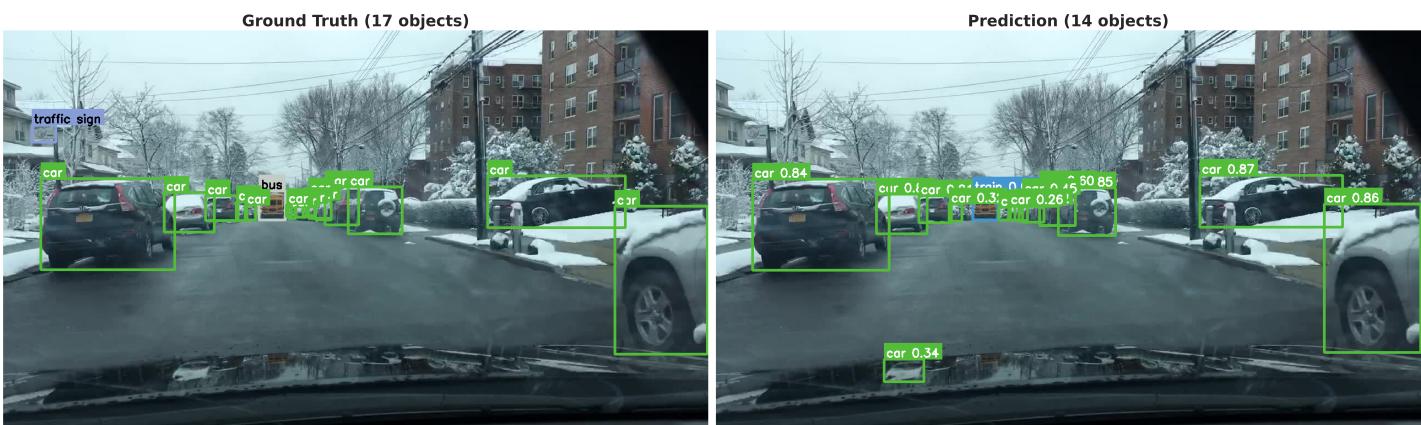
Sample #5: f0d34668-855db02f.jpg

Ground Truth Objects: 12 | Predicted Objects: 1
Weather: clear | Scene: highway | Time of Day: night



Sample #6: d020a2d2-a4b57cb1.jpg

Ground Truth Objects: 17 | Predicted Objects: 14
Weather: snowy | Scene: residential | Time of Day: daytime



Generated by YOLO Testing Pipeline