

YOLO Model Testing Report

Model:	yolov8n
Model Size:	6.2 MB
Parameters:	3.2 M
FLOPs (640x640):	8.86 GFLOPs
Run Name:	yolov8n_testing_20251127_204833
Timestamp:	2025-11-27 21:01:27
Dataset:	bdd100k_yolo_limited - test split
Images Processed:	20000
Total Execution Time	115.31s
IoU Threshold:	0.5

Inference Performance

Metric	Value
Average Inference Time	1.51 ms per image
FPS (Frames Per Second)	663.76

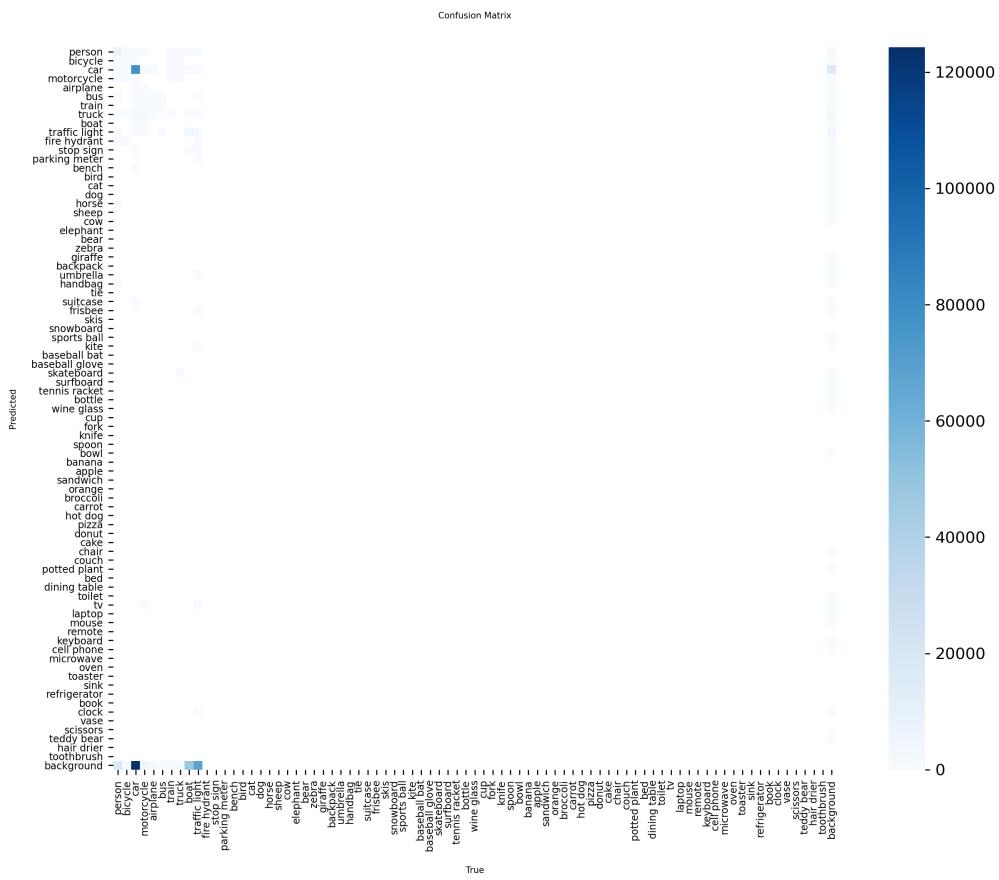
Note: Inference time includes preprocessing, model inference, and postprocessing.

Overall Accuracy Metrics

Metric	Value
Precision	0.2336
Recall	0.6904
F1-Score	0.3491
mAP@0.5	0.0924
mAP@0.5:0.95	0.0479
True Positives	85908
False Positives	281819
False Negatives	38517

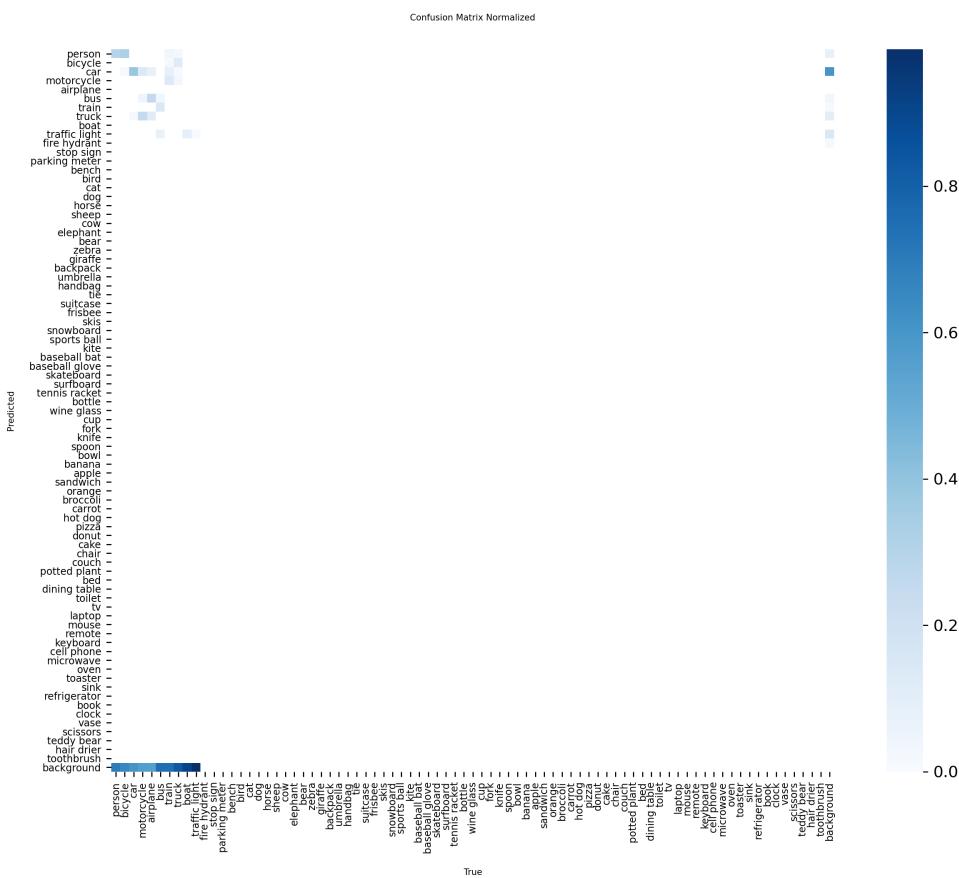
Confusion Matrices

Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

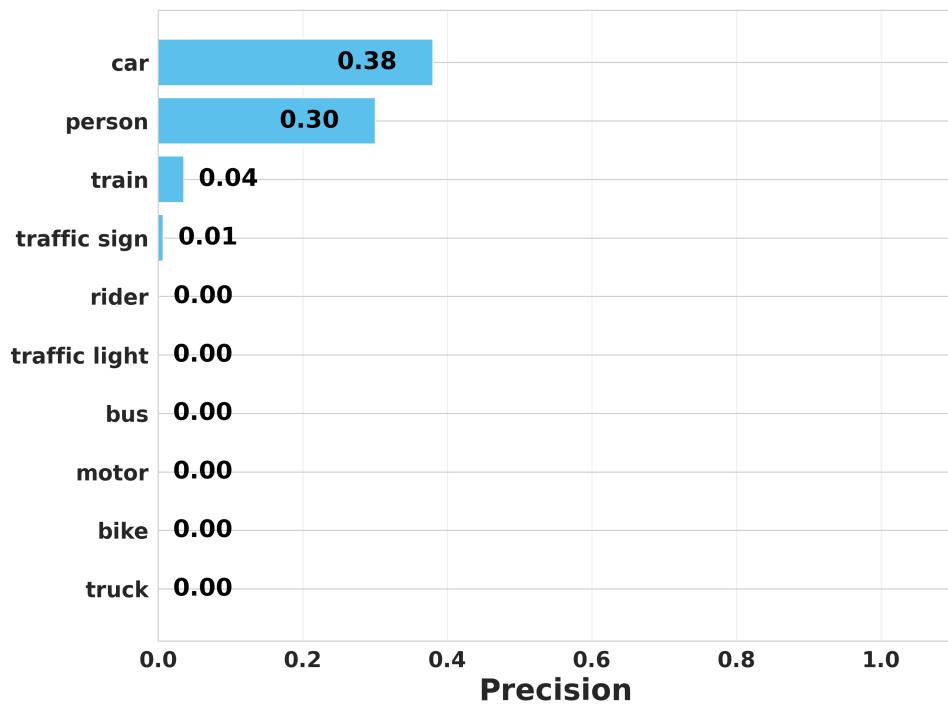
Confusion Matrix (Normalized)



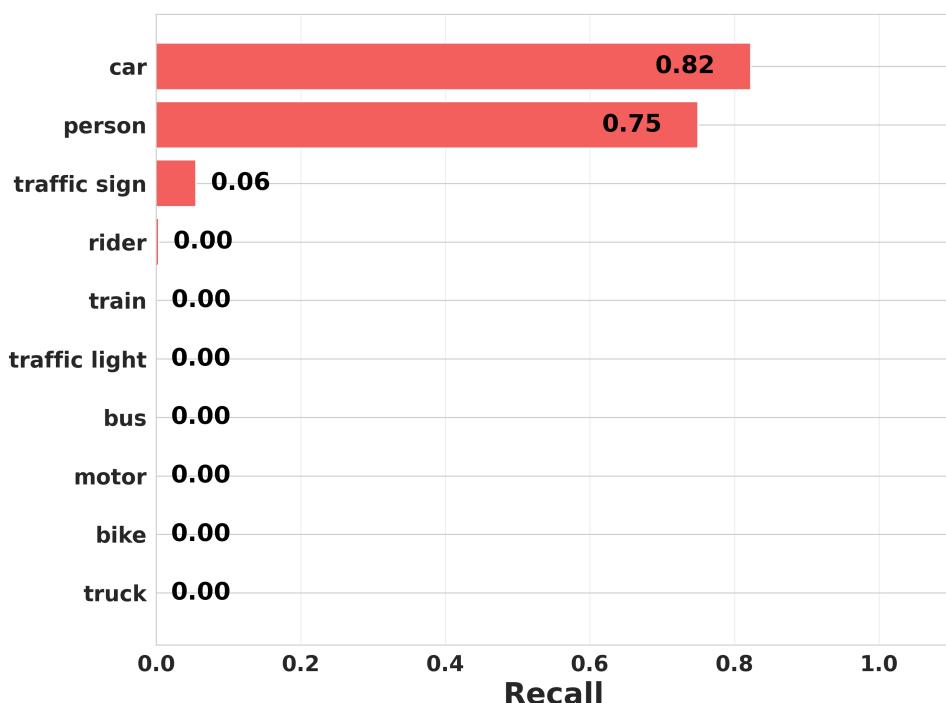
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

Performance Visualizations

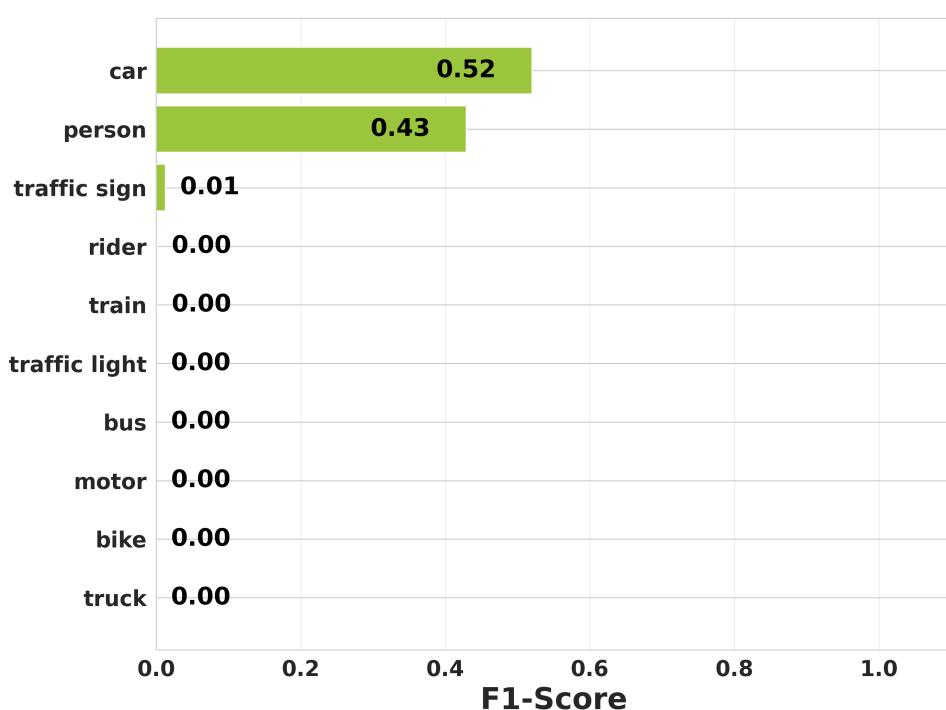
Precision by Class



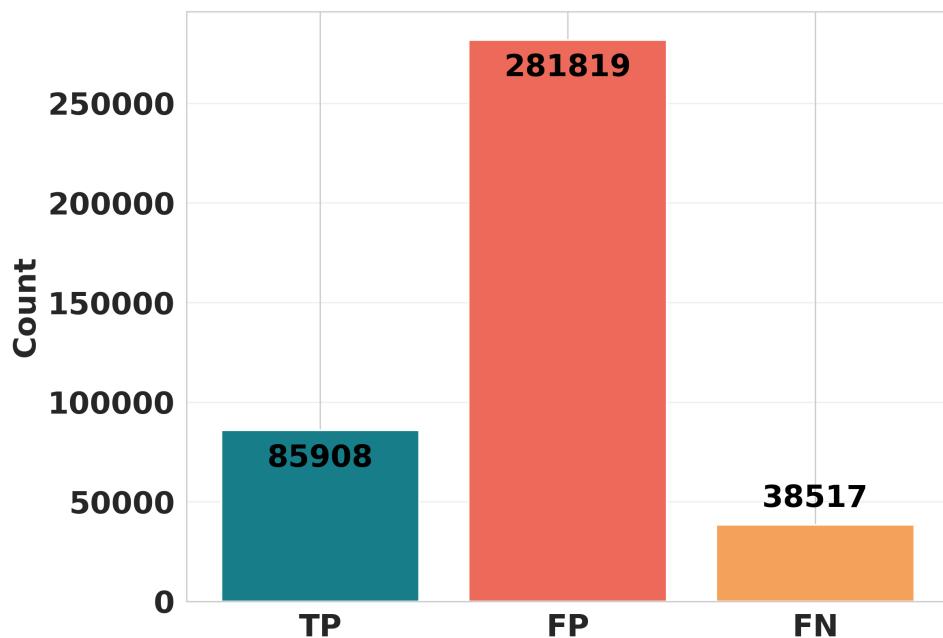
Recall by Class



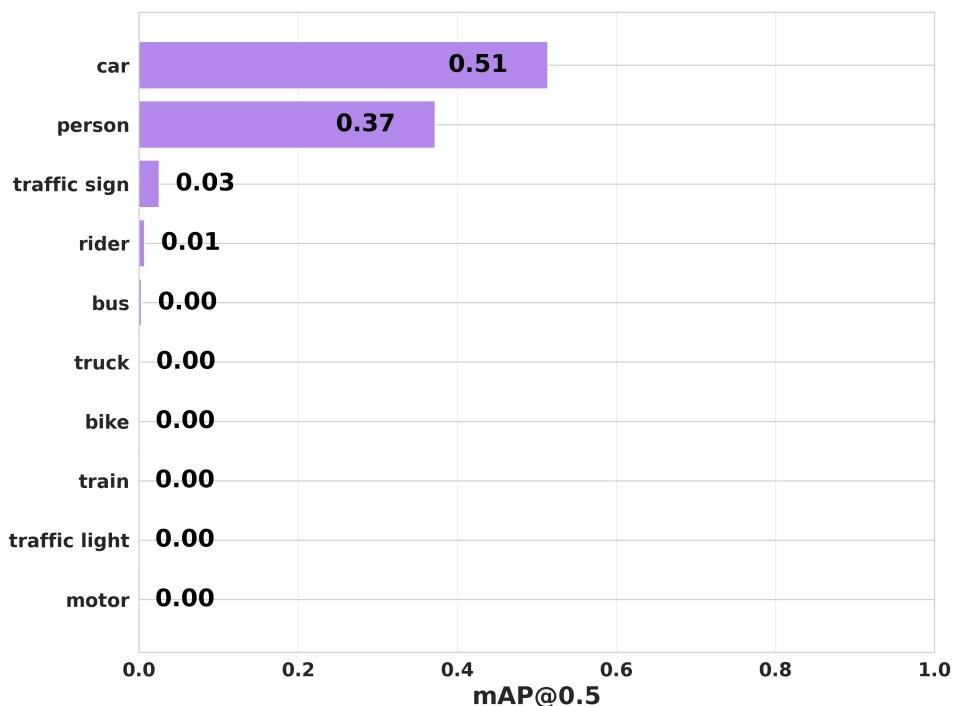
F1-Score by Class



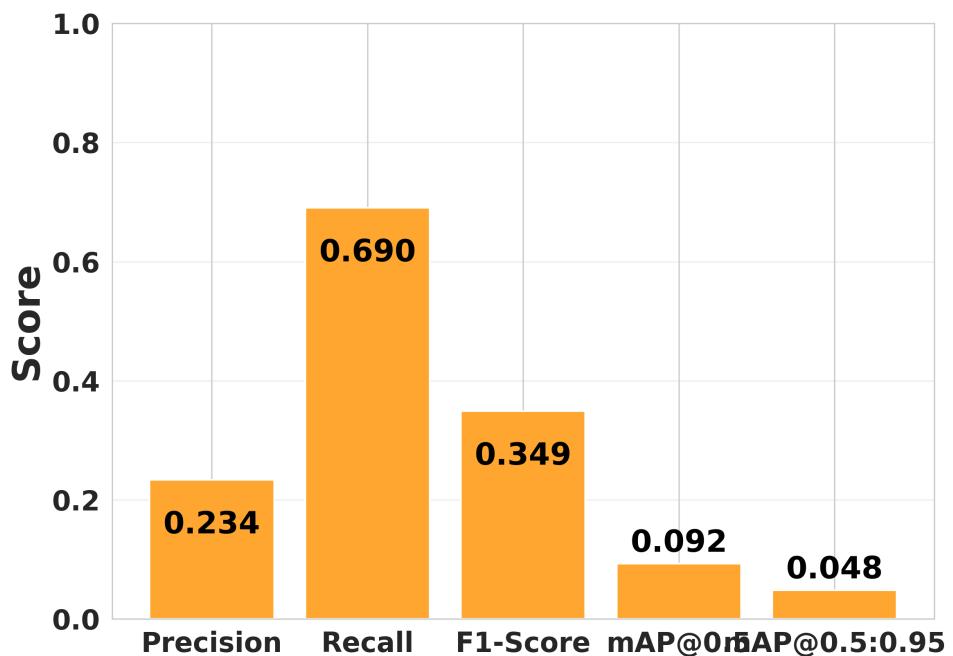
Overall Detection Outcomes



mAP@0.5 by Class



Overall Metrics



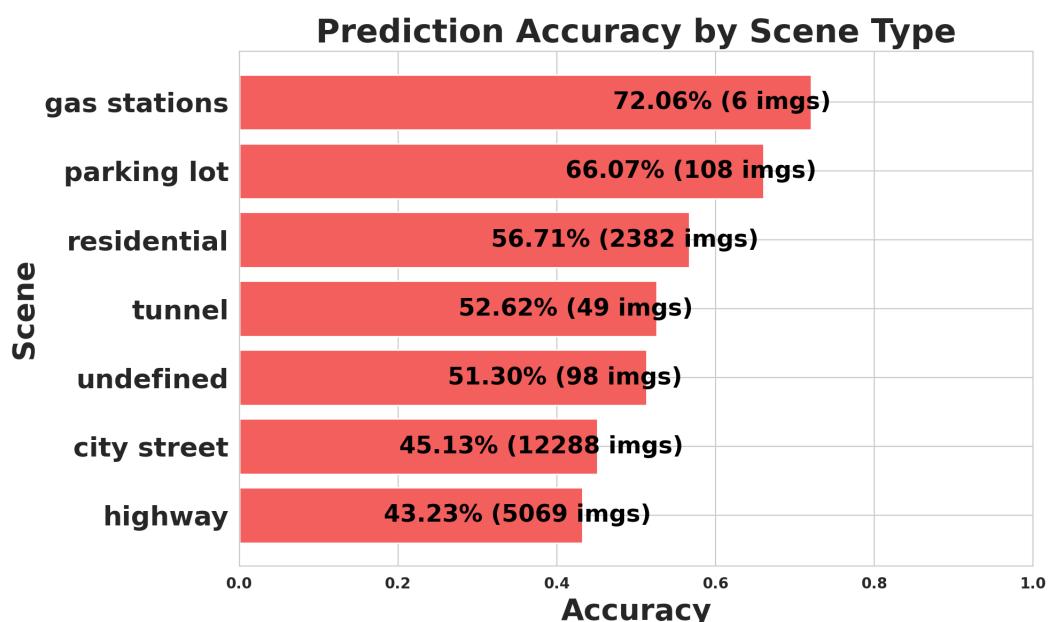
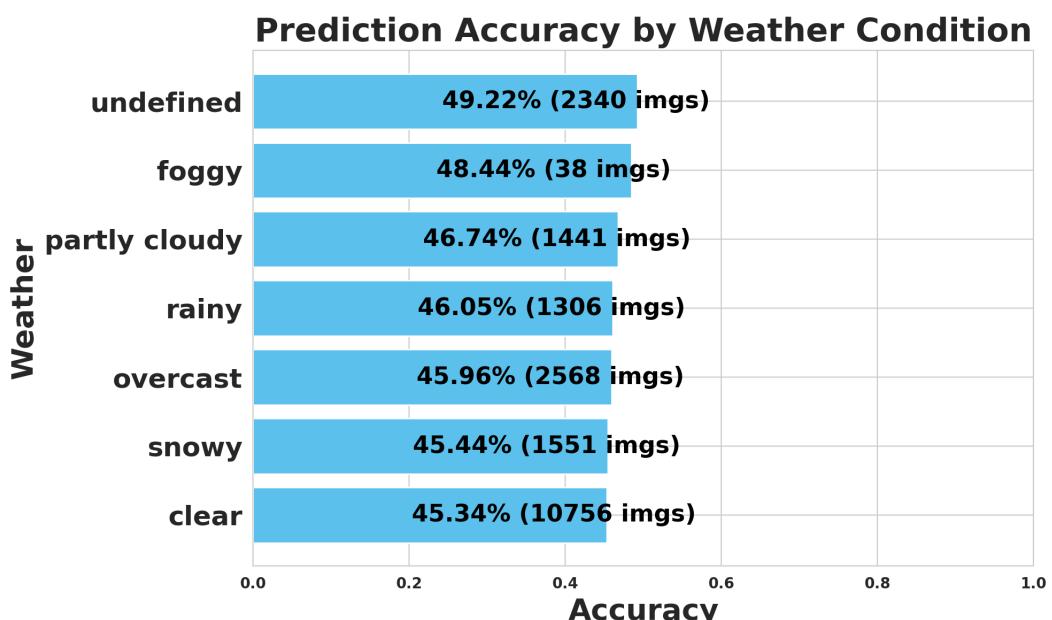
Per-Class Performance

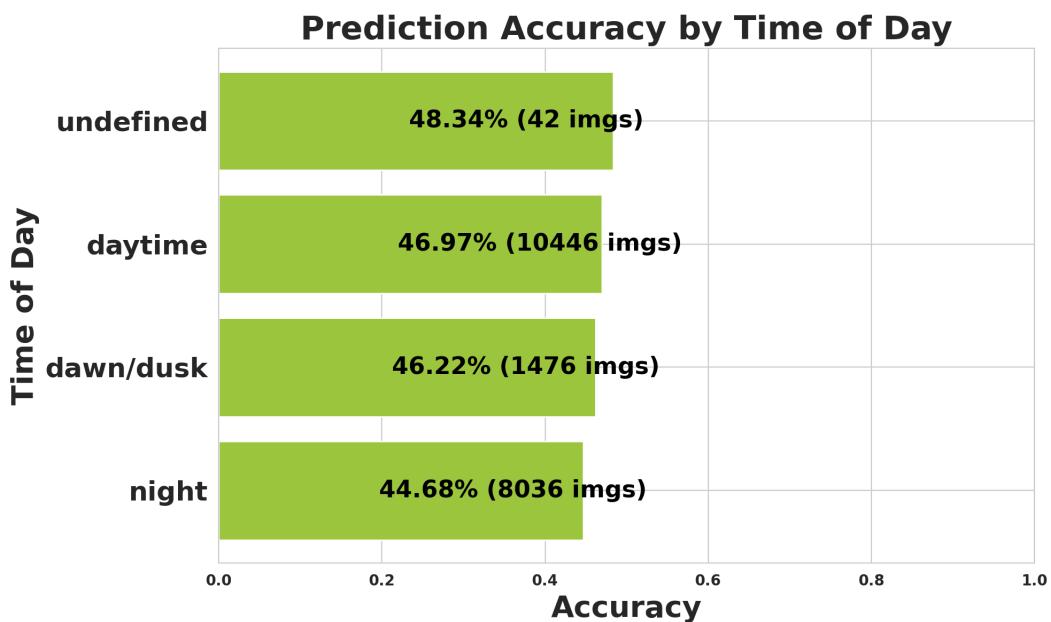
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	7413	17237	2479	0.3007	0.7494	0.4292	0.3725
rider	1	1293	299	0.0008	0.0033	0.0013	0.0072
car	78001	127148	16796	0.3802	0.8228	0.5201	0.5140
truck	0	8704	230	0.0000	0.0000	0.0000	0.0013
bus	0	3217	35	0.0000	0.0000	0.0000	0.0031
train	1	27	2406	0.0357	0.0004	0.0008	0.0002
motor	0	841	237	0.0000	0.0000	0.0000	0.0001
bike	0	1998	7566	0.0000	0.0000	0.0000	0.0002
traffic light	0	52840	27	0.0000	0.0000	0.0000	0.0001
traffic sign	492	68514	8442	0.0071	0.0551	0.0126	0.0255

Comprehensive Failure Analysis

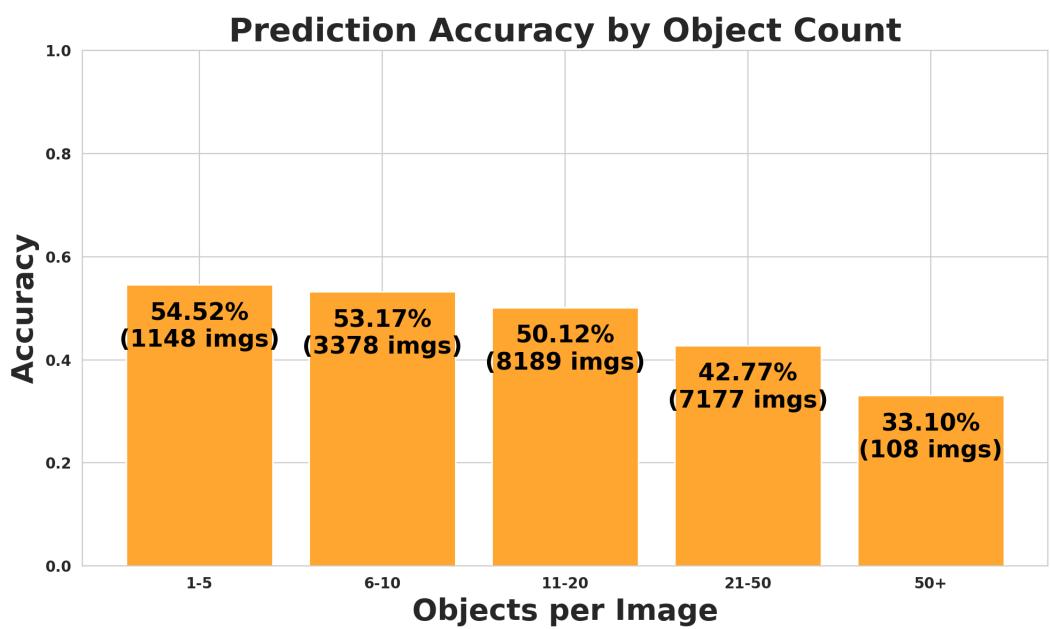
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	169488
Overall Accuracy	46.09%

Accuracy by Environmental Attributes

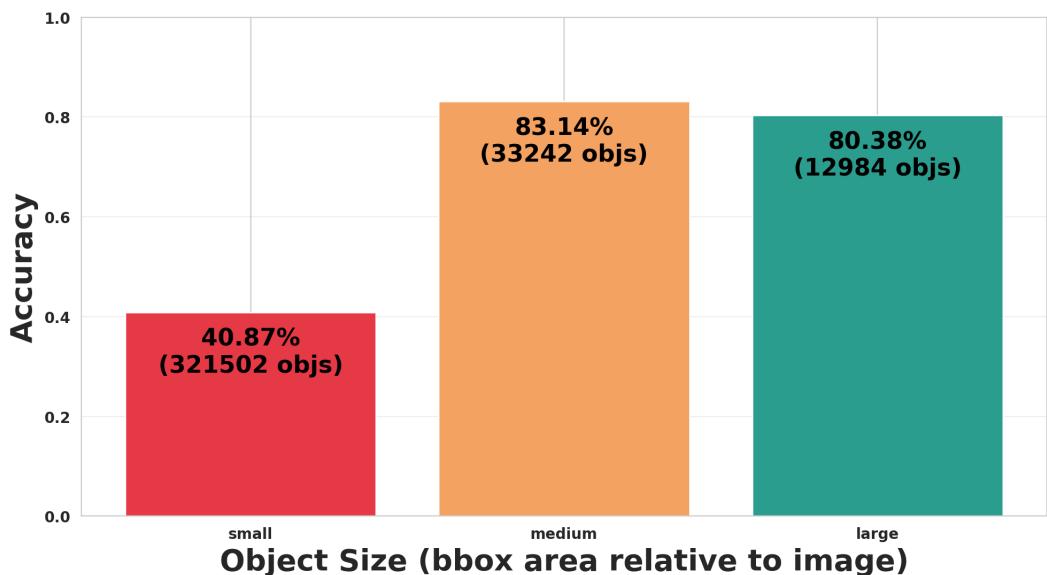




Accuracy by Object Characteristics

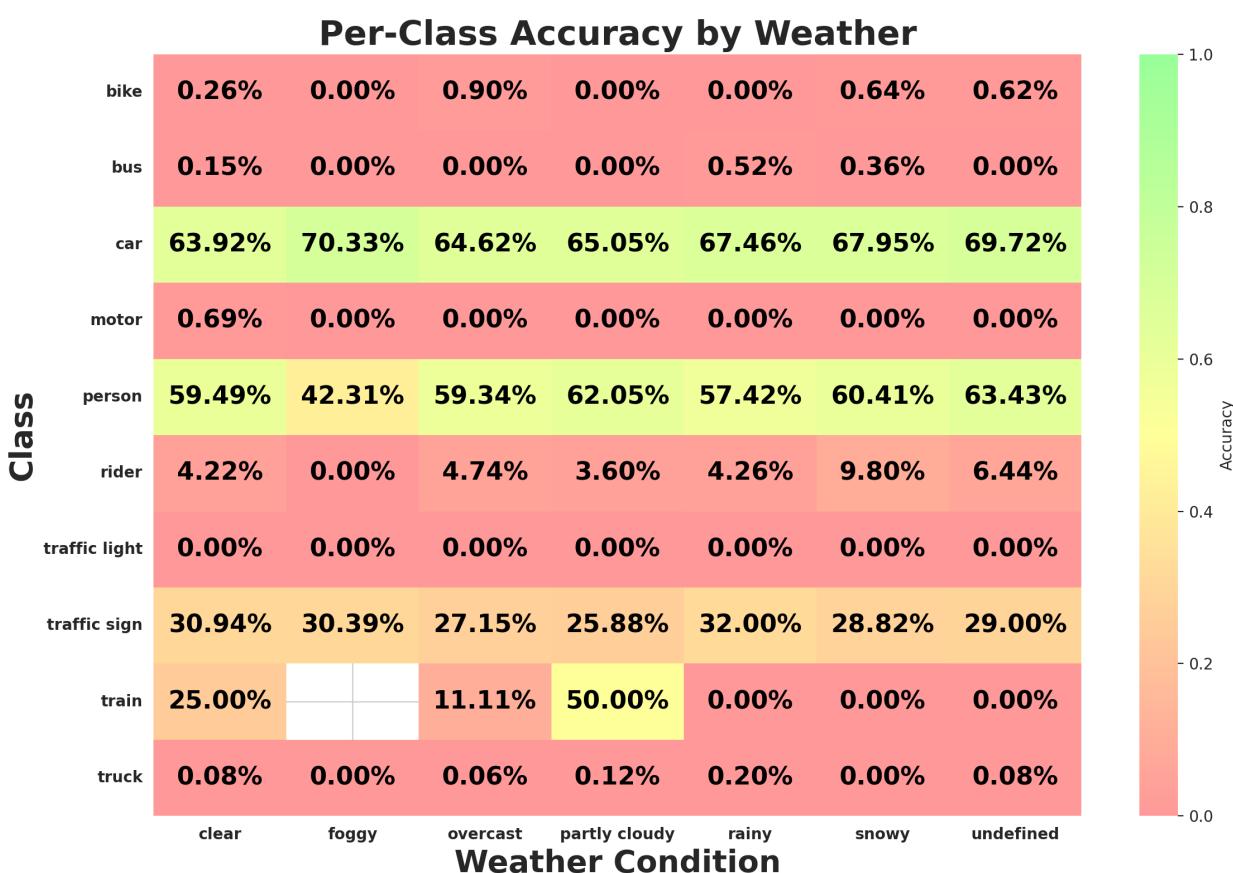
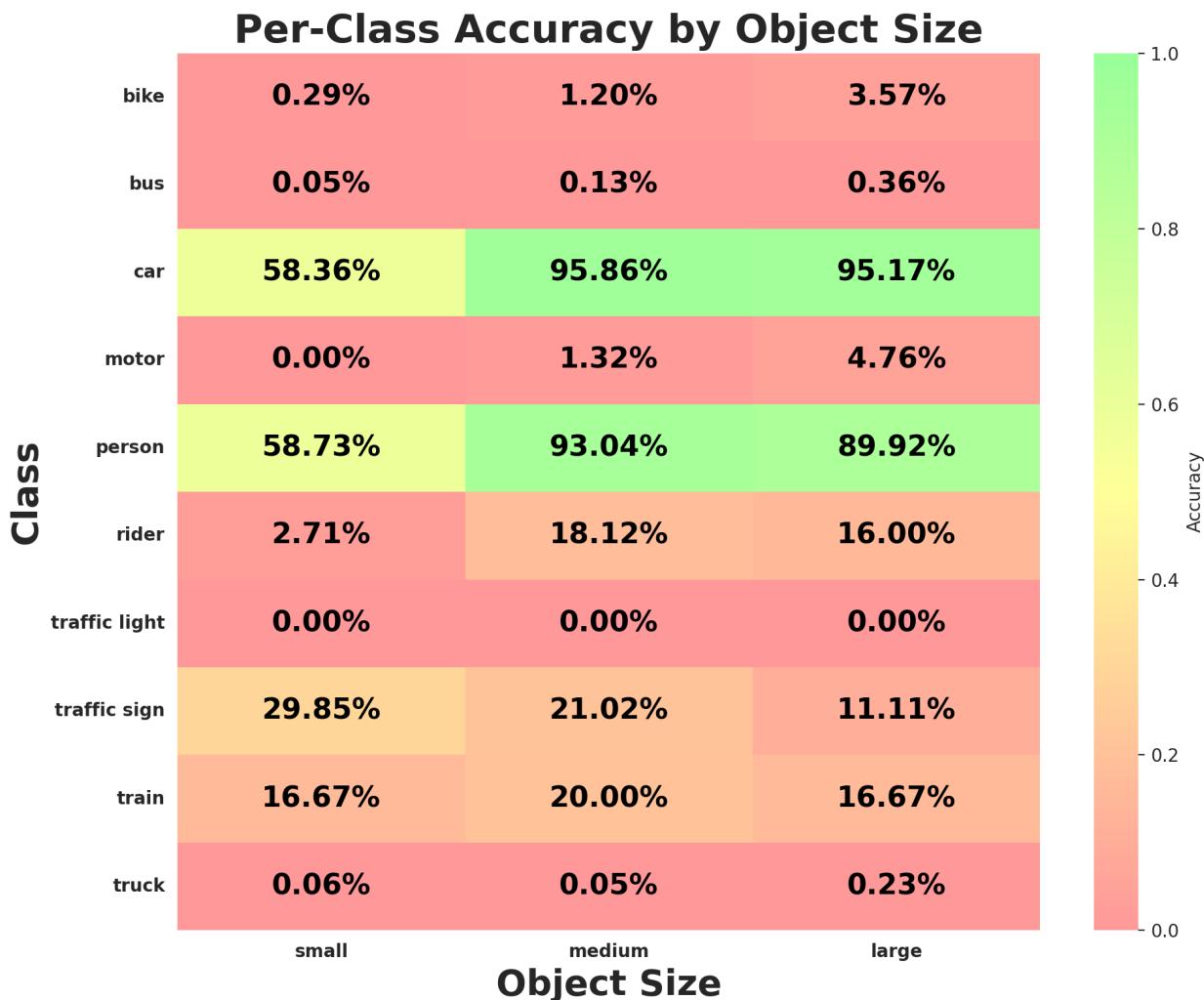


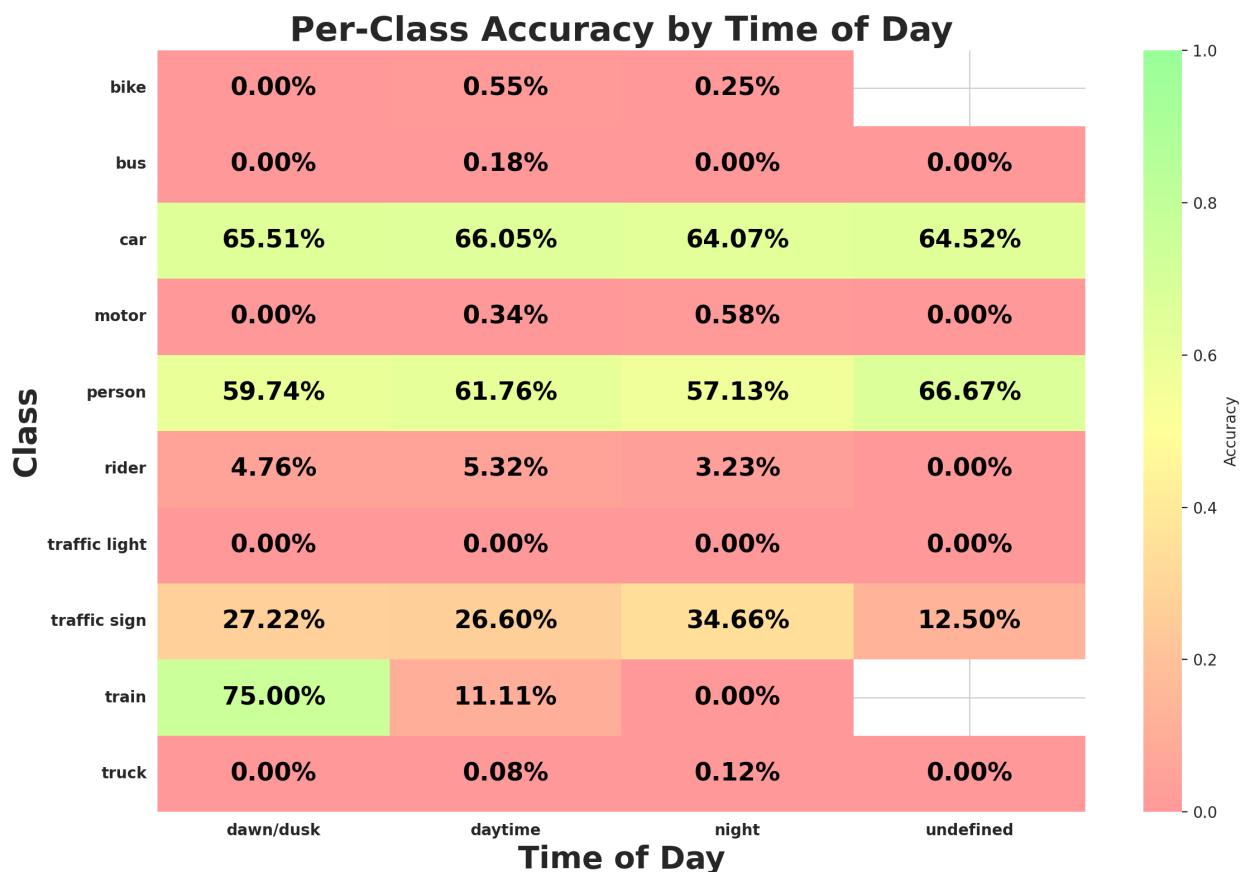
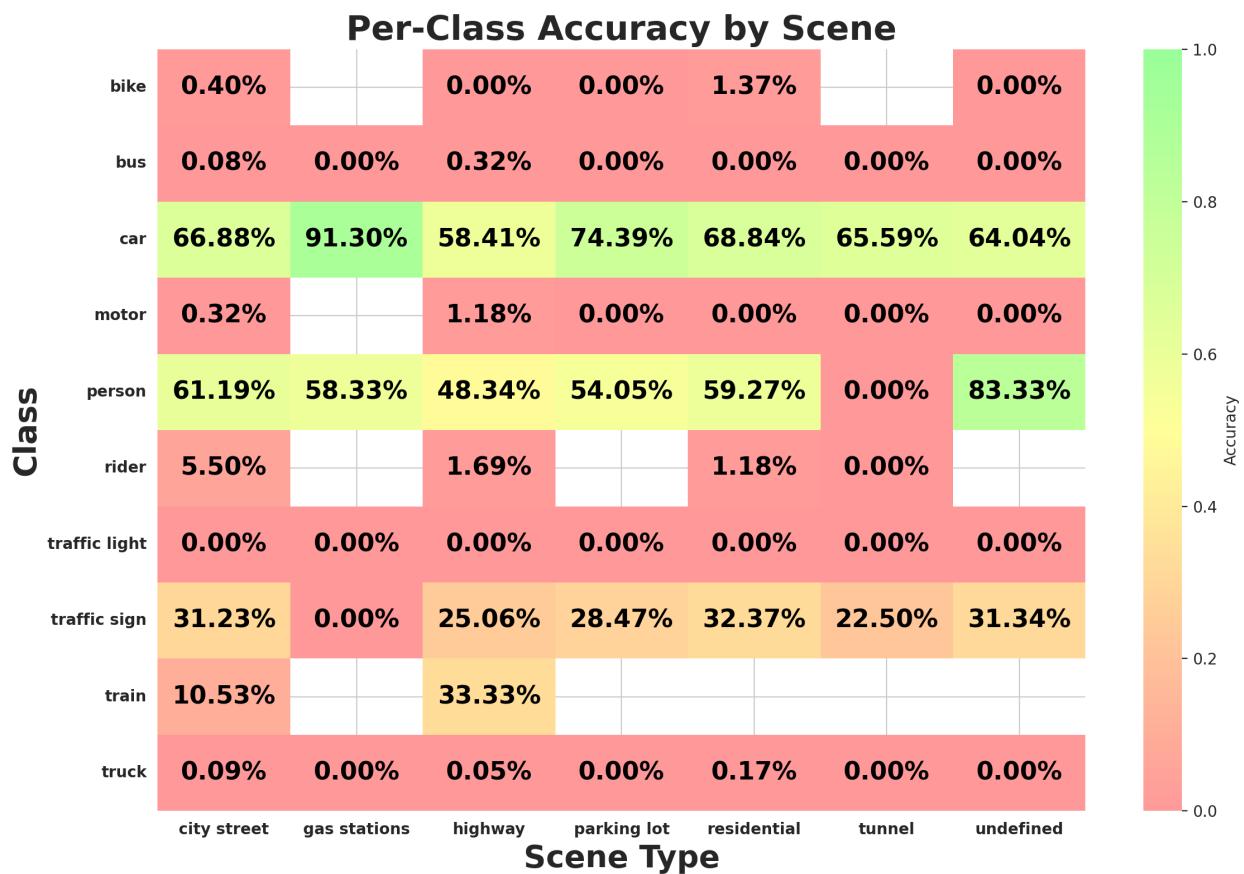
Prediction Accuracy by Object Size



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

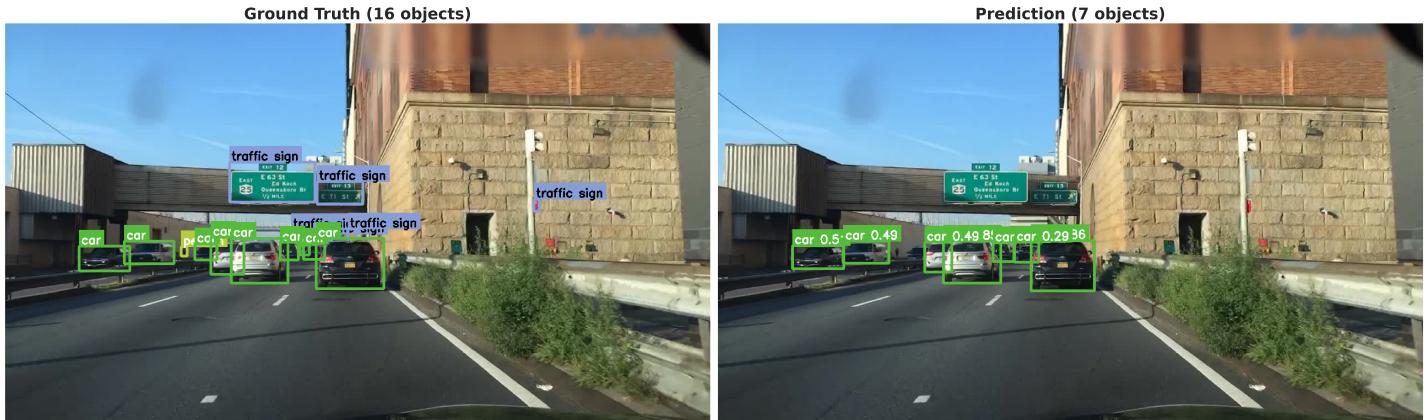




Sample Predictions: Ground Truth vs Model

Sample #1: ea6aaaf5a-e6cf85a2.jpg

Ground Truth Objects: 16 | Predicted Objects: 7
Weather: clear | Scene: highway | Time of Day: daytime



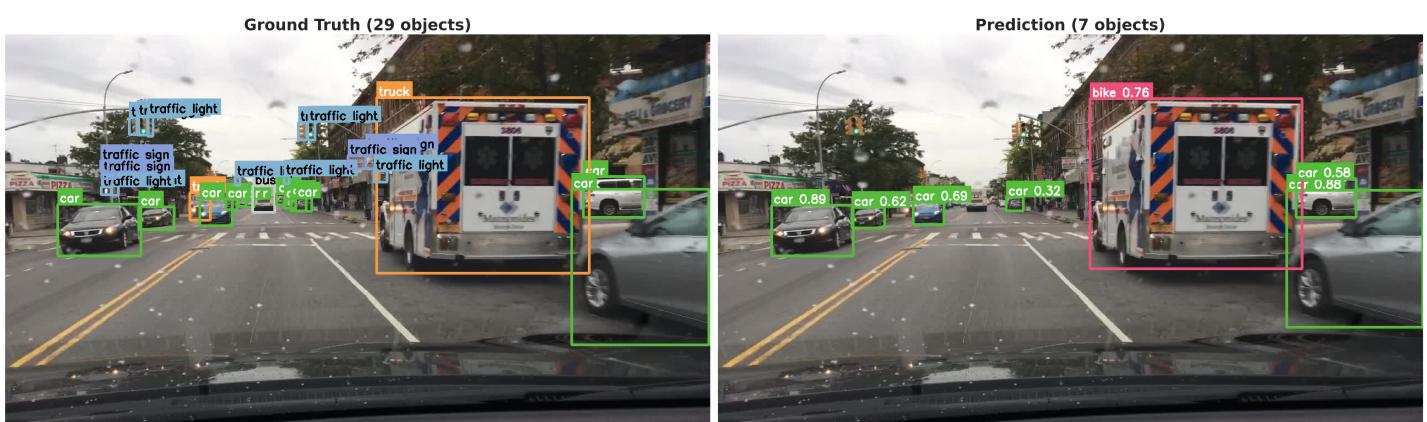
Sample #2: d834f383-ce645189.jpg

Ground Truth Objects: 11 | Predicted Objects: 4
Weather: undefined | Scene: city street | Time of Day: daytime



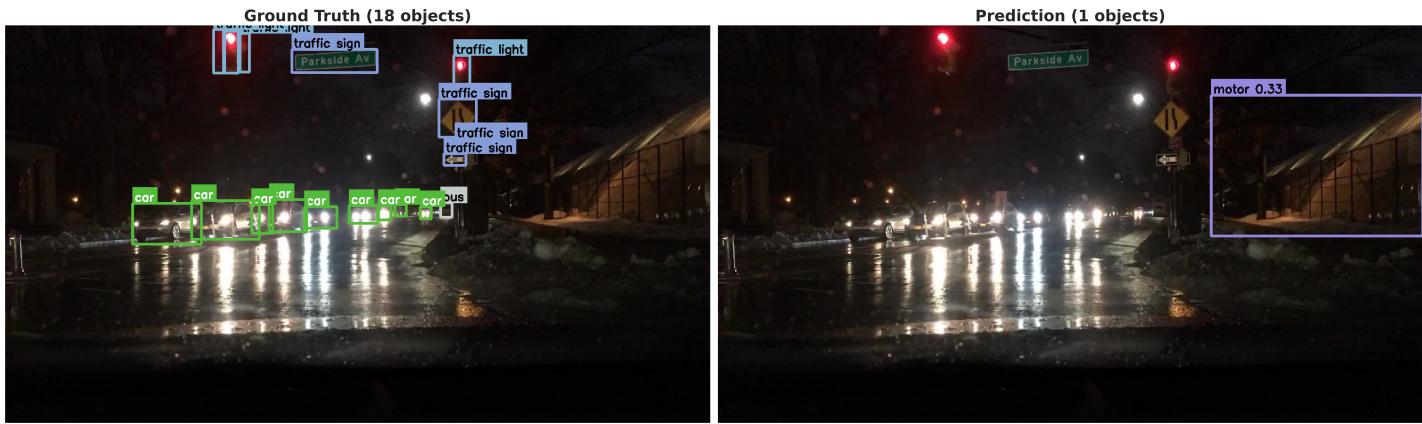
Sample #3: f05854c2-bc1a1eef.jpg

Ground Truth Objects: 29 | Predicted Objects: 7
Weather: rainy | Scene: city street | Time of Day: daytime



Sample #4: eed5b8dc-dab9d0de.jpg

Ground Truth Objects: 18 | Predicted Objects: 1
Weather: rainy | Scene: city street | Time of Day: night



Sample #5: fc0253c5-114c02c7.jpg

Ground Truth Objects: 24 | Predicted Objects: 4
Weather: clear | Scene: residential | Time of Day: daytime



Sample #6: e3502587-fab6ee91.jpg

Ground Truth Objects: 5 | Predicted Objects: 4
Weather: rainy | Scene: residential | Time of Day: daytime



Generated by YOLO Testing Pipeline