

YOLO Model Testing Report

Model:	yolov10x
Model Size:	61.4 MB
Parameters:	31.8 M
FLOPs (640x640):	171.85 GFLOPs
Run Name:	yolov10x_testing_20251127_235254
Timestamp:	2025-11-28 00:09:25
Dataset:	bdd100k_yolo_limited - test split
Images Processed:	20000
Total Execution Time	141.43s
IoU Threshold:	0.5

Inference Performance

Metric	Value
Average Inference Time	3.61 ms per image
FPS (Frames Per Second)	276.86

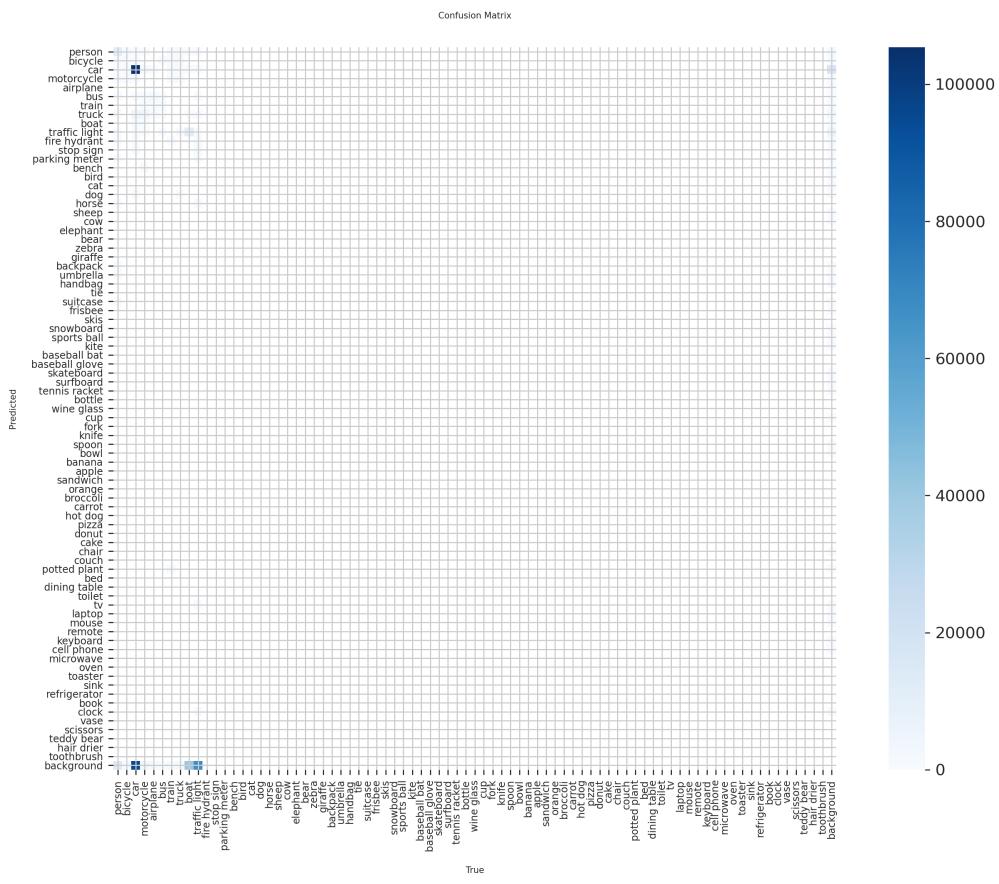
Note: Inference time includes preprocessing, model inference, and postprocessing.

Overall Accuracy Metrics

Metric	Value
Precision	0.3176
Recall	0.6947
F1-Score	0.4359
mAP@0.5	0.1259
mAP@0.5:0.95	0.0709
True Positives	116783
False Positives	250944
False Negatives	51316

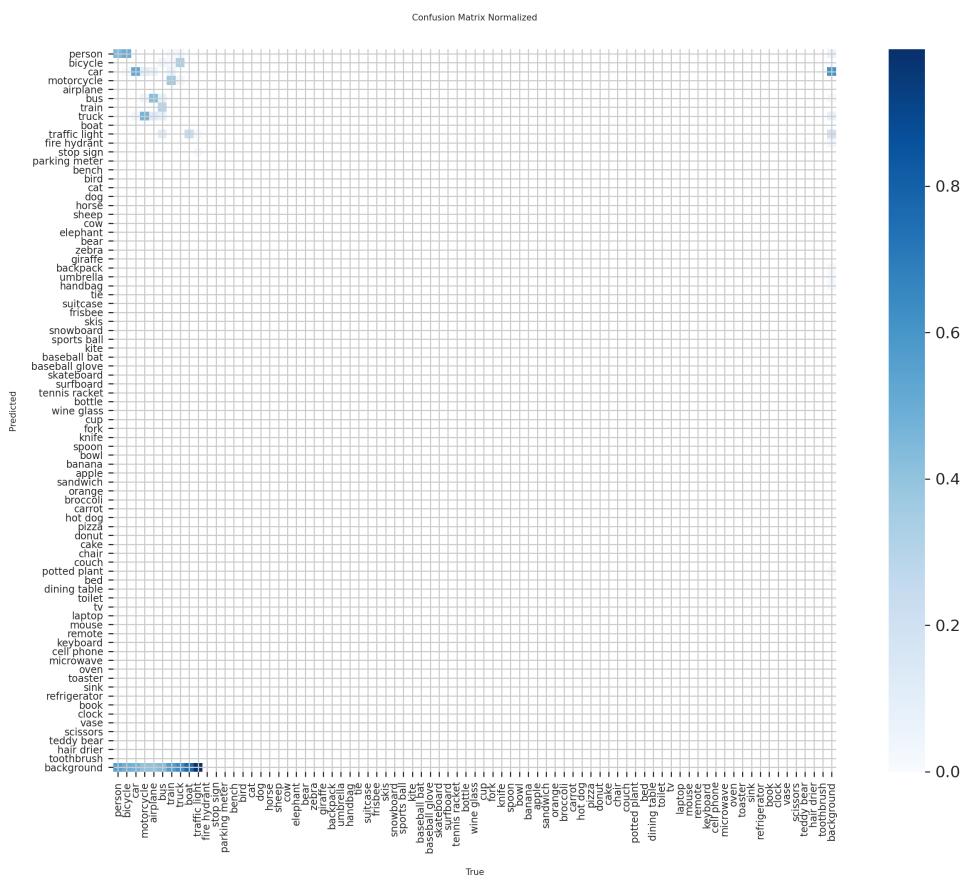
Confusion Matrices

Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

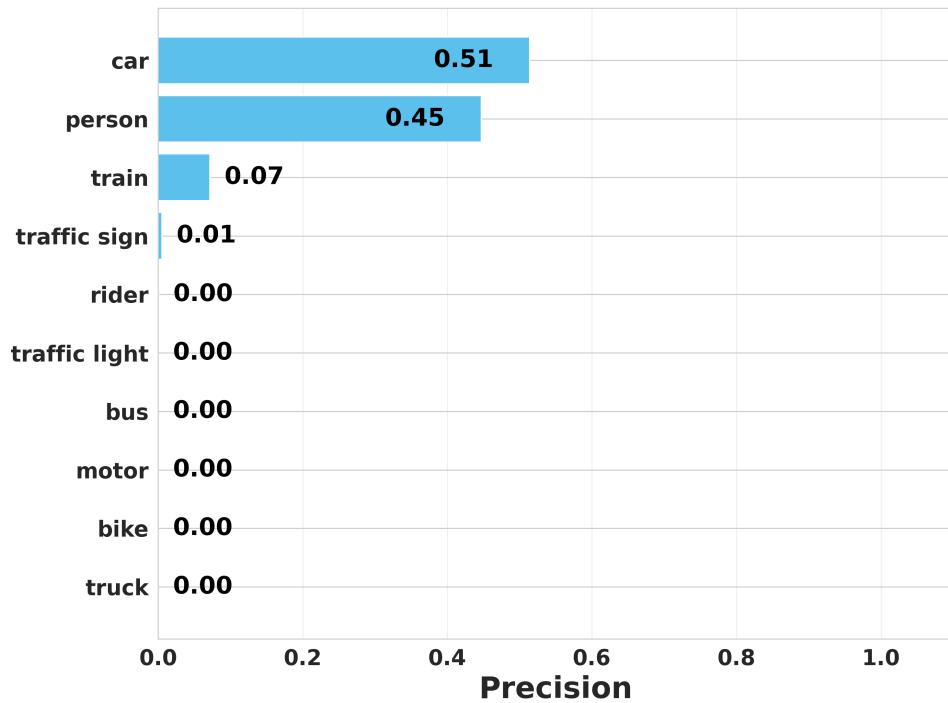
Confusion Matrix (Normalized)



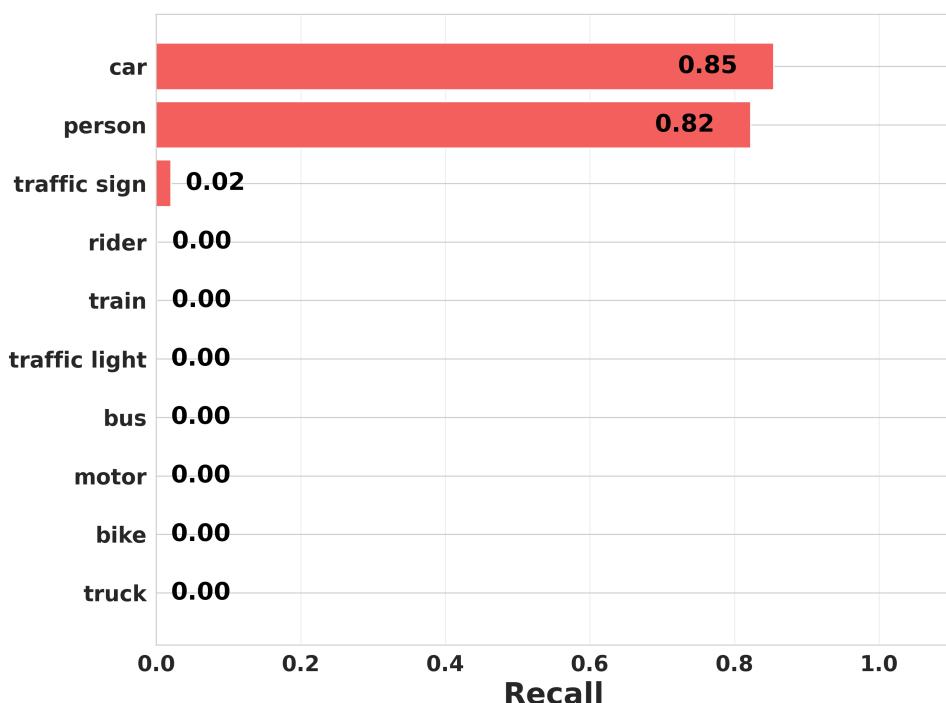
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

Performance Visualizations

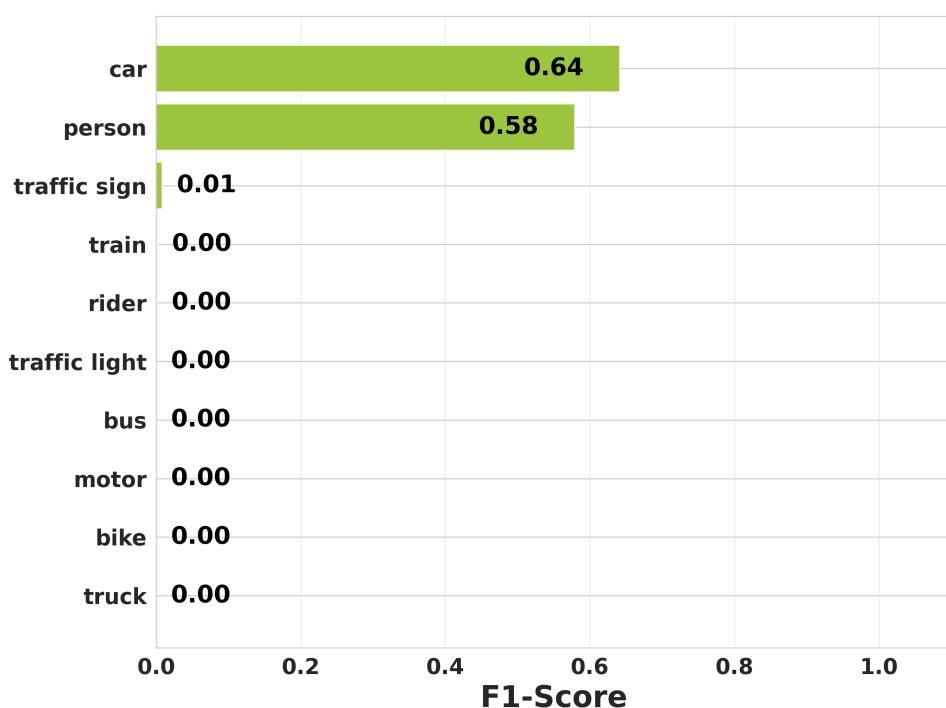
Precision by Class



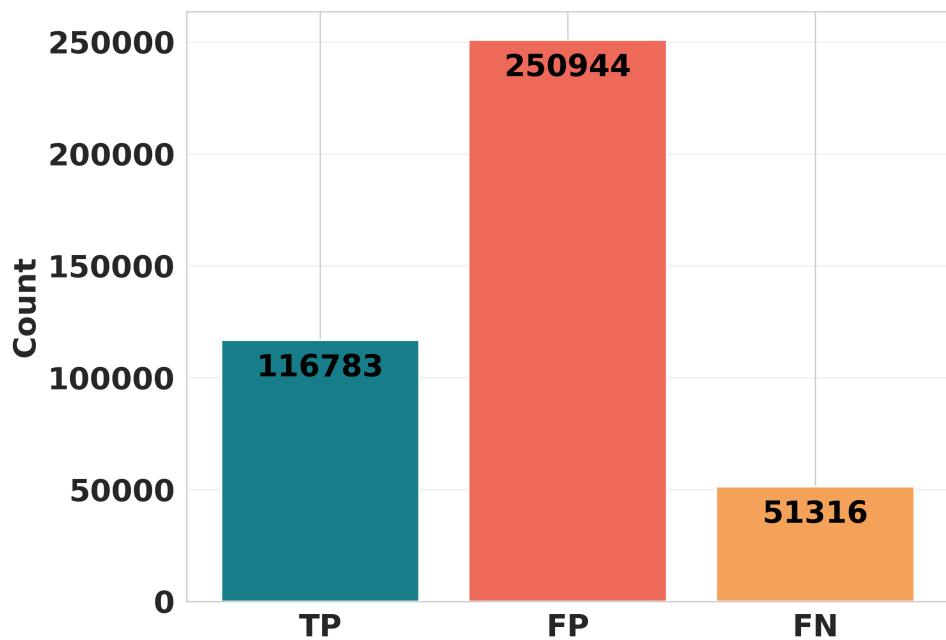
Recall by Class



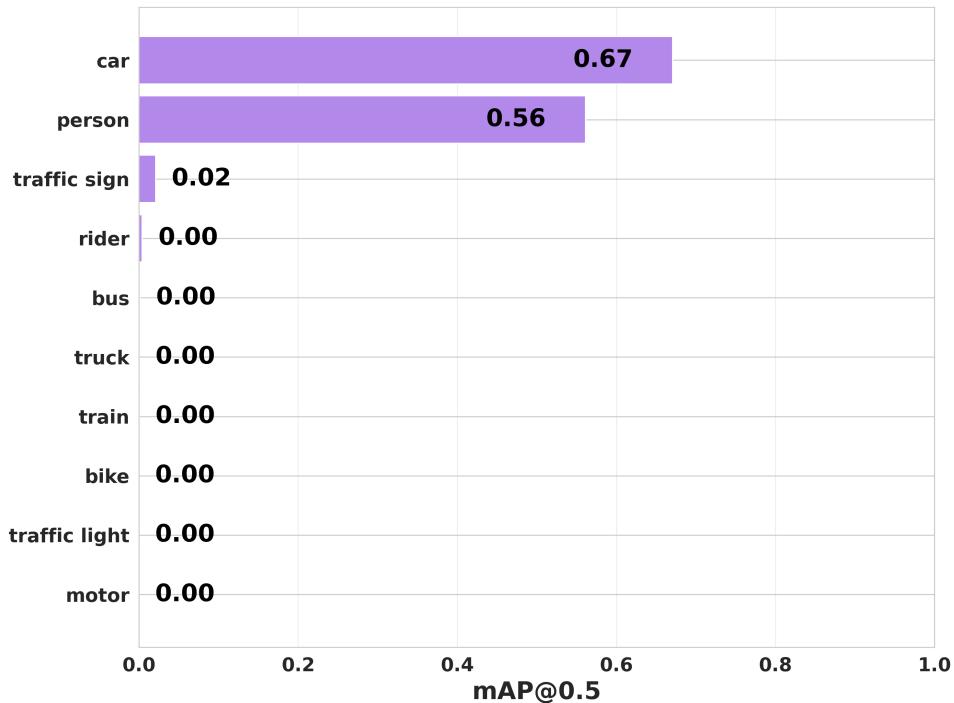
F1-Score by Class



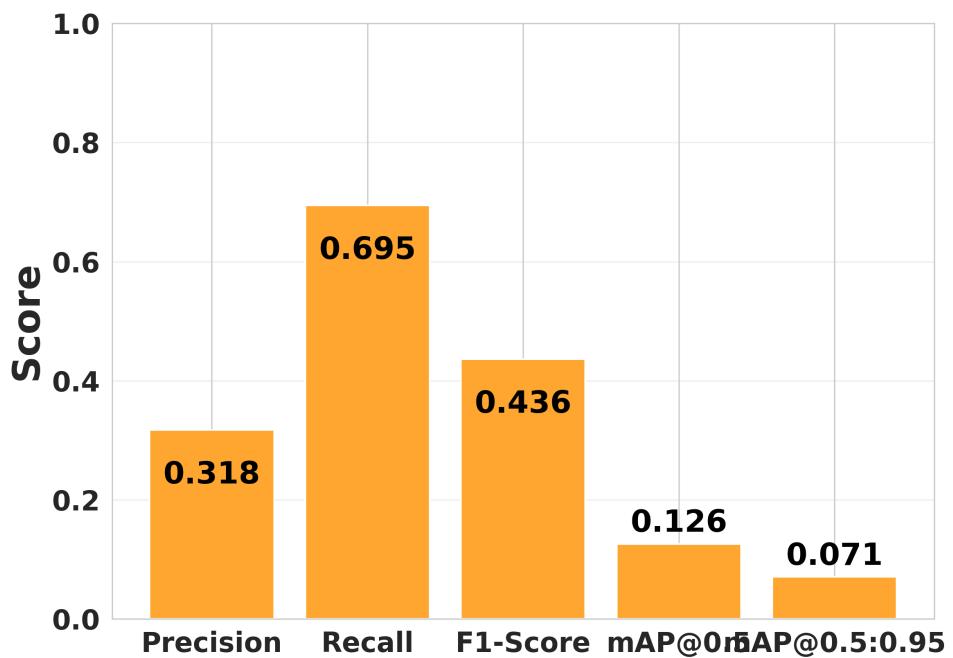
Overall Detection Outcomes



mAP@0.5 by Class



Overall Metrics



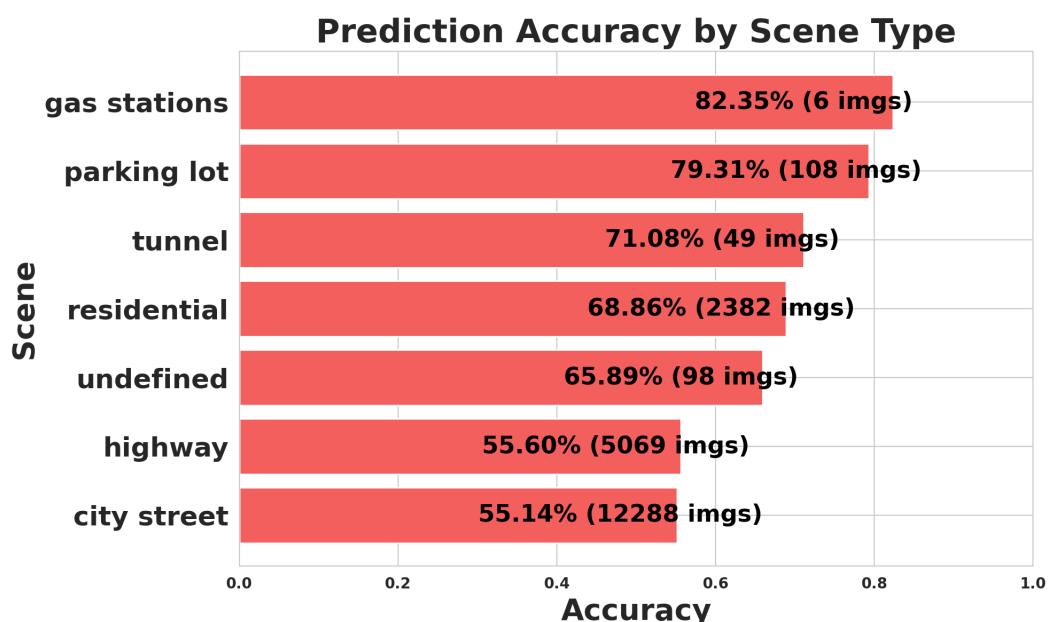
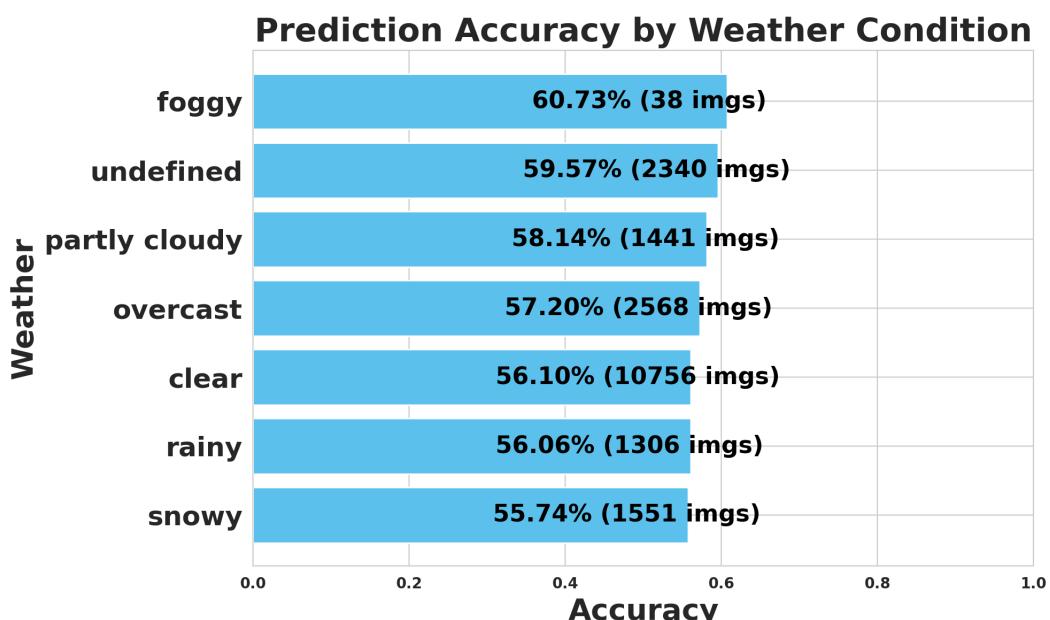
Per-Class Performance

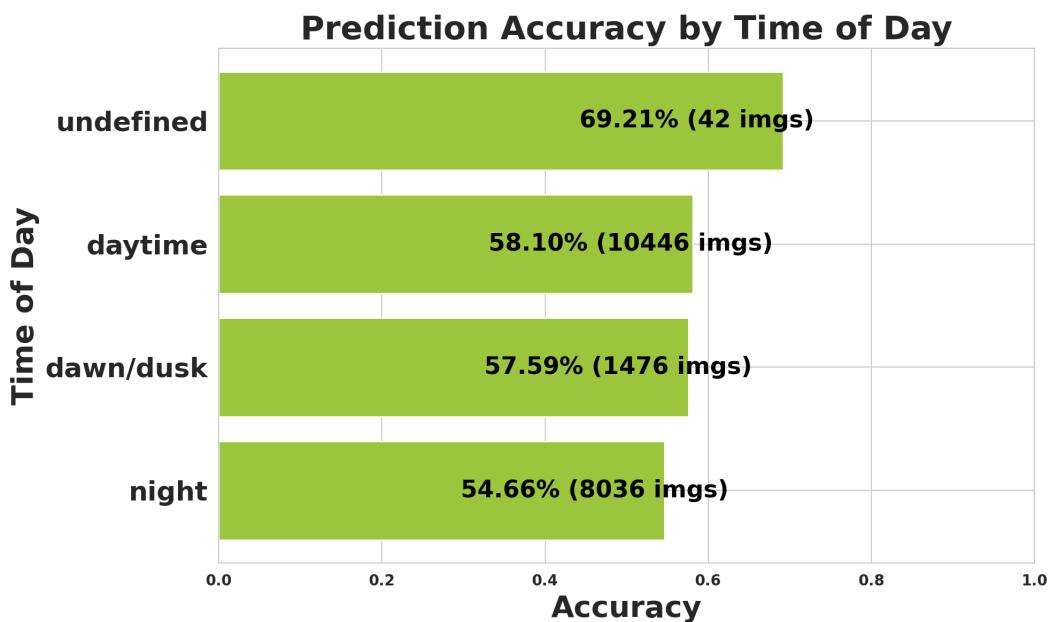
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	11015	13635	2379	0.4469	0.8224	0.5791	0.5613
rider	1	1293	781	0.0008	0.0013	0.0010	0.0040
car	105403	99746	17936	0.5138	0.8546	0.6417	0.6710
truck	0	8704	444	0.0000	0.0000	0.0000	0.0004
bus	0	3217	10	0.0000	0.0000	0.0000	0.0012
train	2	26	2346	0.0714	0.0009	0.0017	0.0003
motor	0	841	85	0.0000	0.0000	0.0000	0.0000
bike	0	1998	9993	0.0000	0.0000	0.0000	0.0000
traffic light	0	52840	10	0.0000	0.0000	0.0000	0.0000
traffic sign	362	68644	17332	0.0052	0.0205	0.0084	0.0209

Comprehensive Failure Analysis

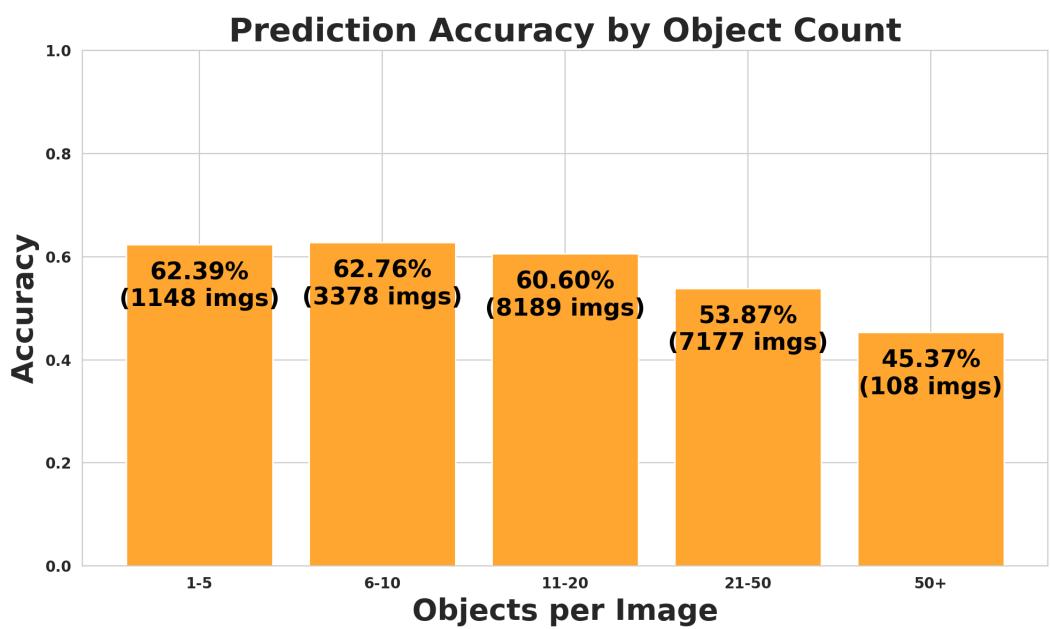
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	209010
Overall Accuracy	56.84%

Accuracy by Environmental Attributes

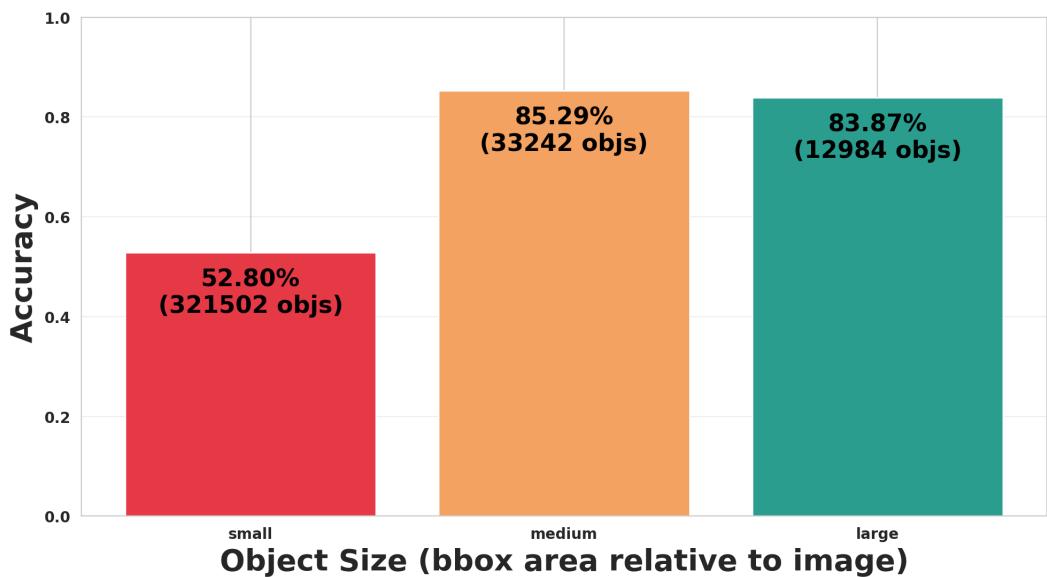




Accuracy by Object Characteristics

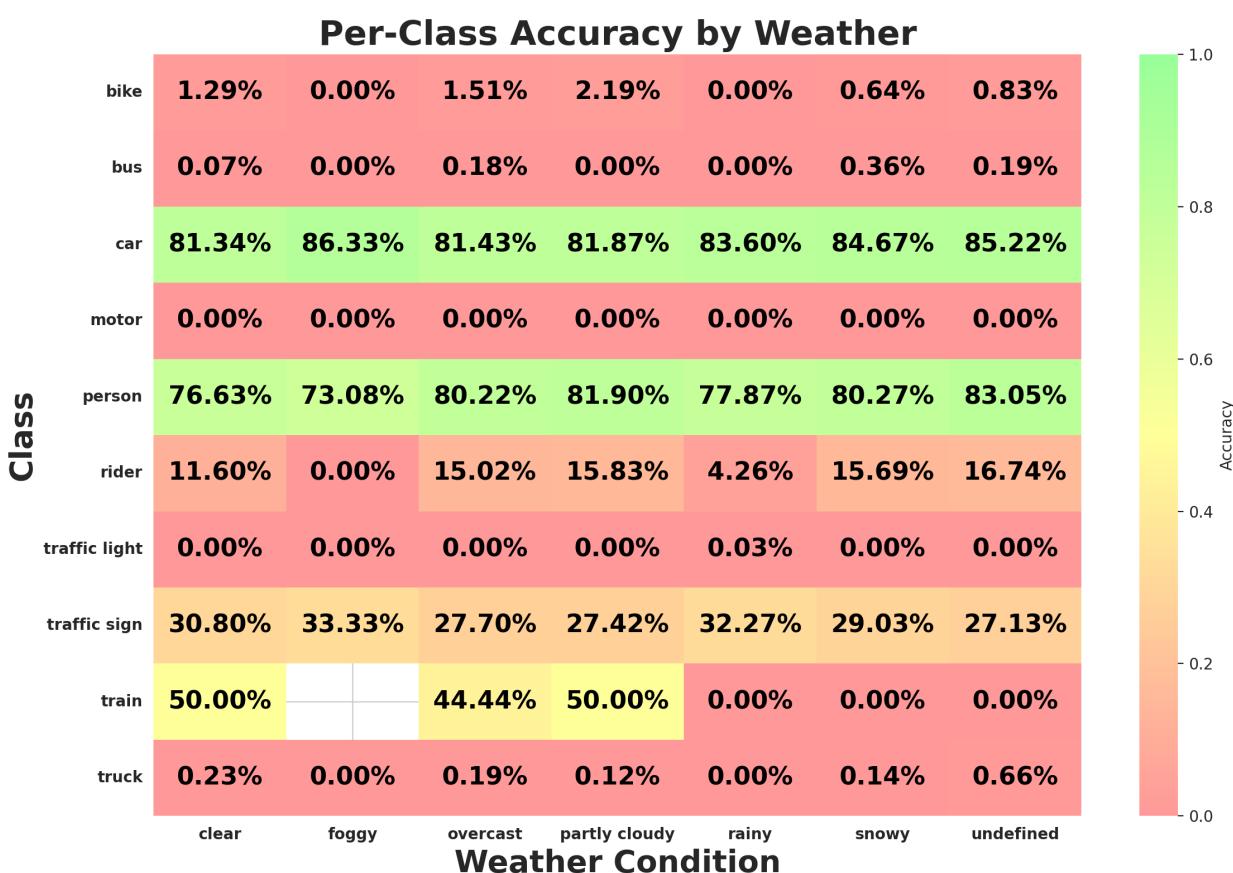
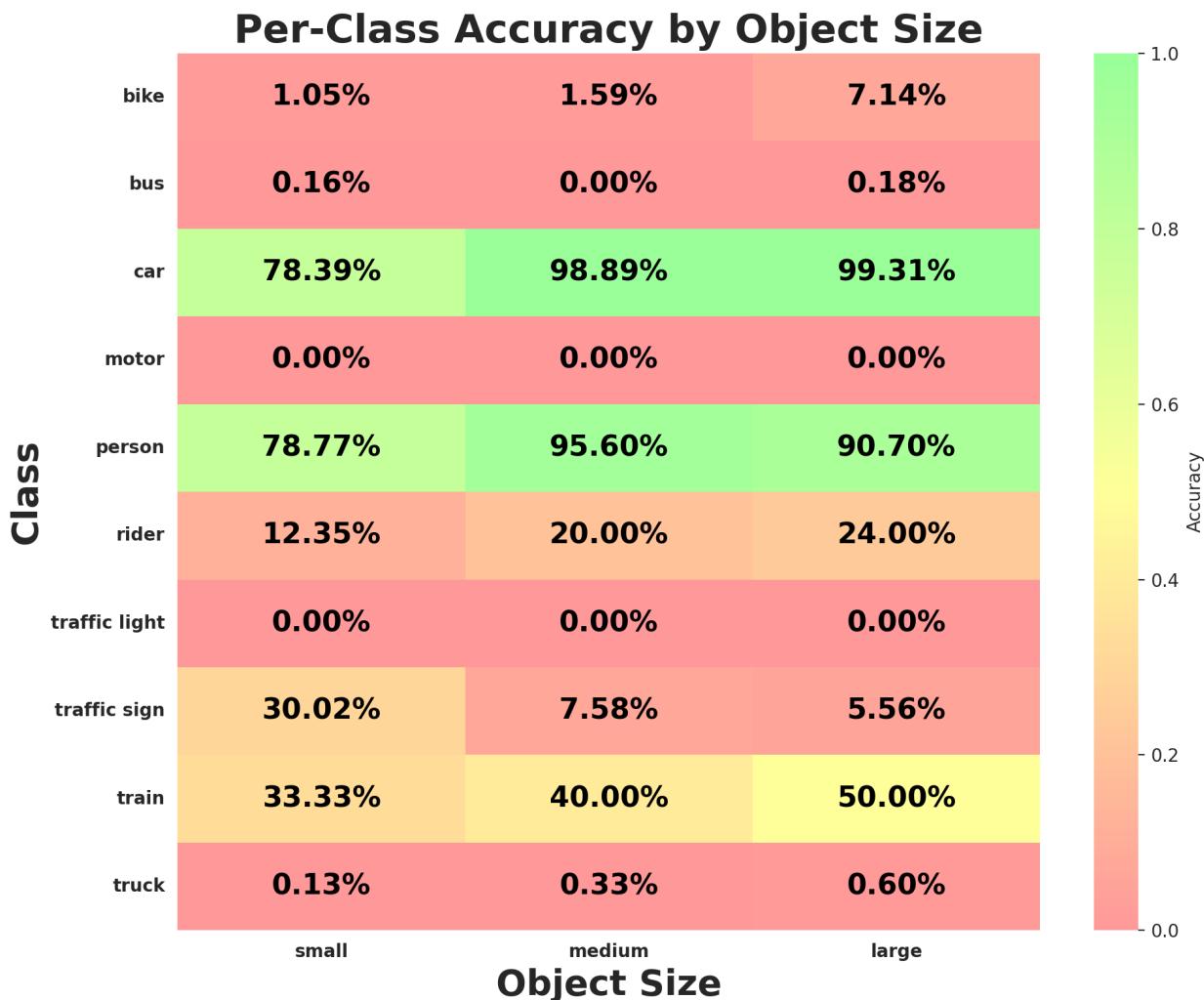


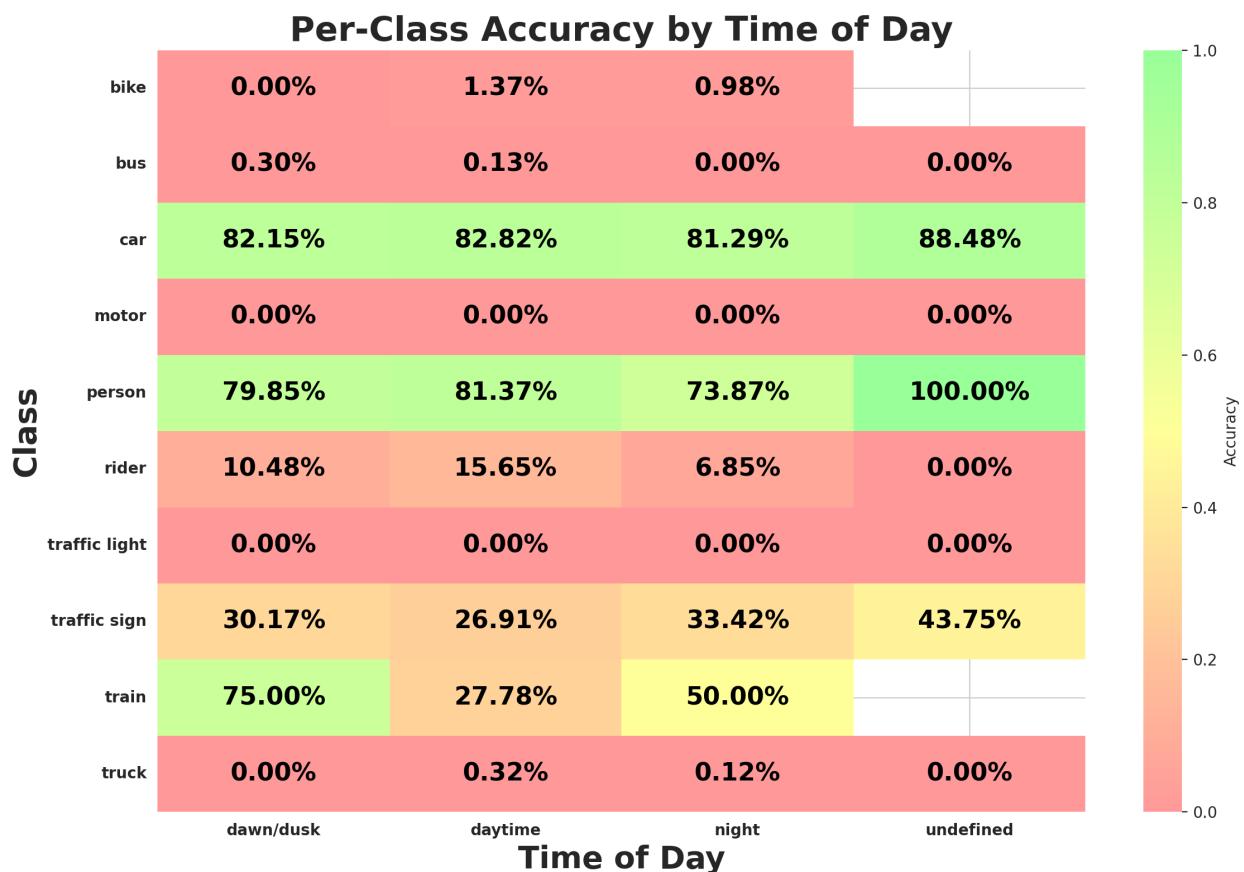
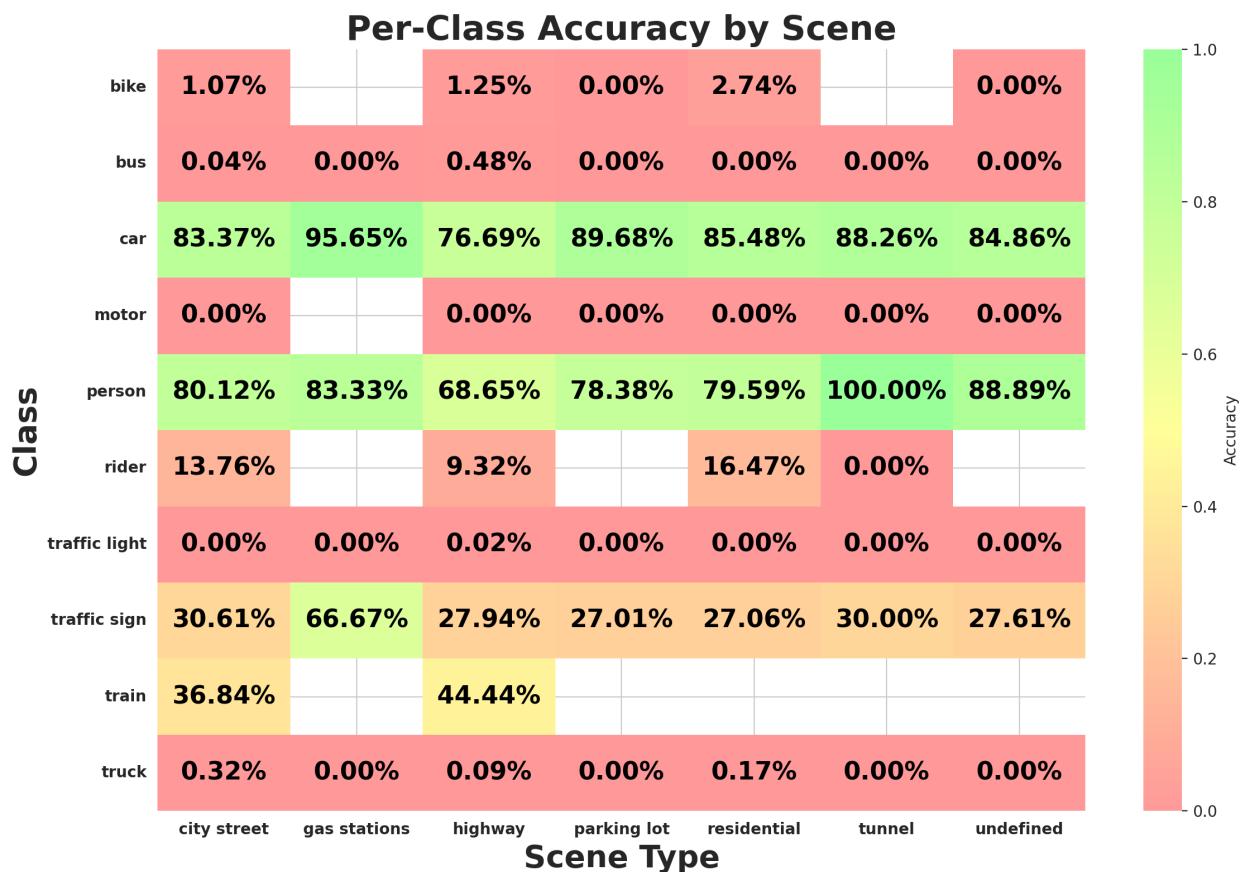
Prediction Accuracy by Object Size



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

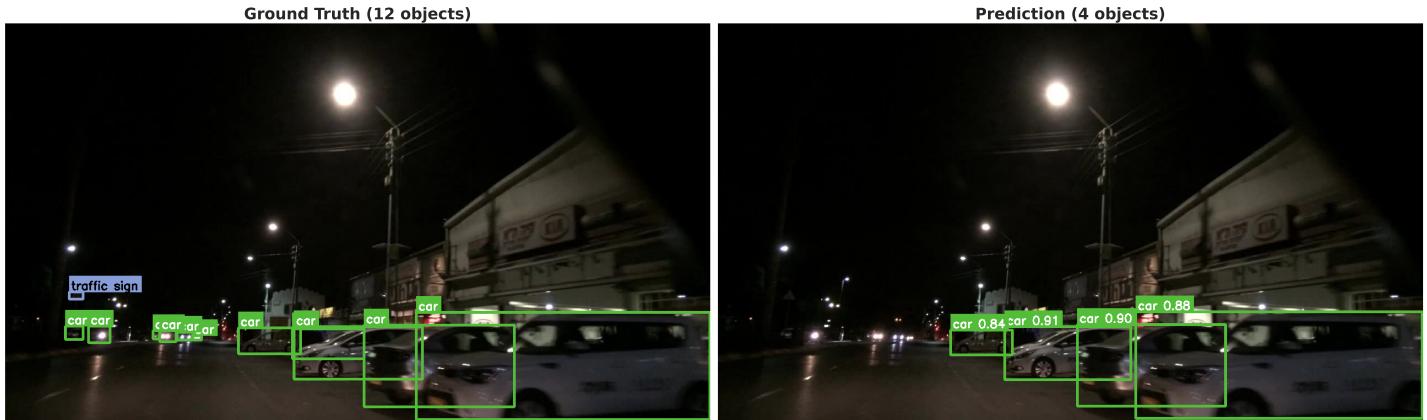




Sample Predictions: Ground Truth vs Model

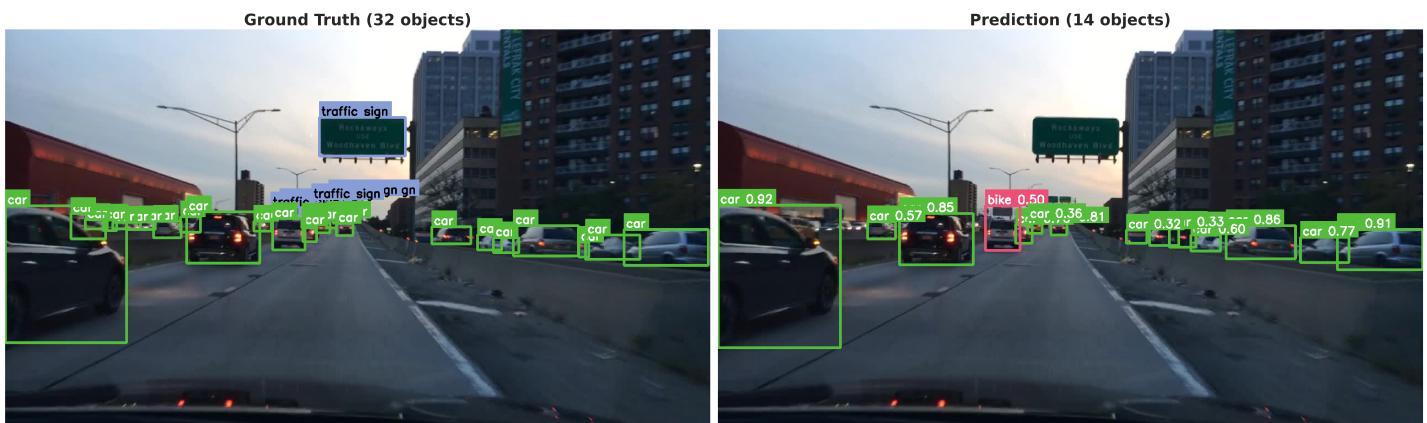
Sample #1: ccb9e408-acf2f3e6.jpg

Ground Truth Objects: 12 | Predicted Objects: 4
Weather: clear | Scene: city street | Time of Day: night



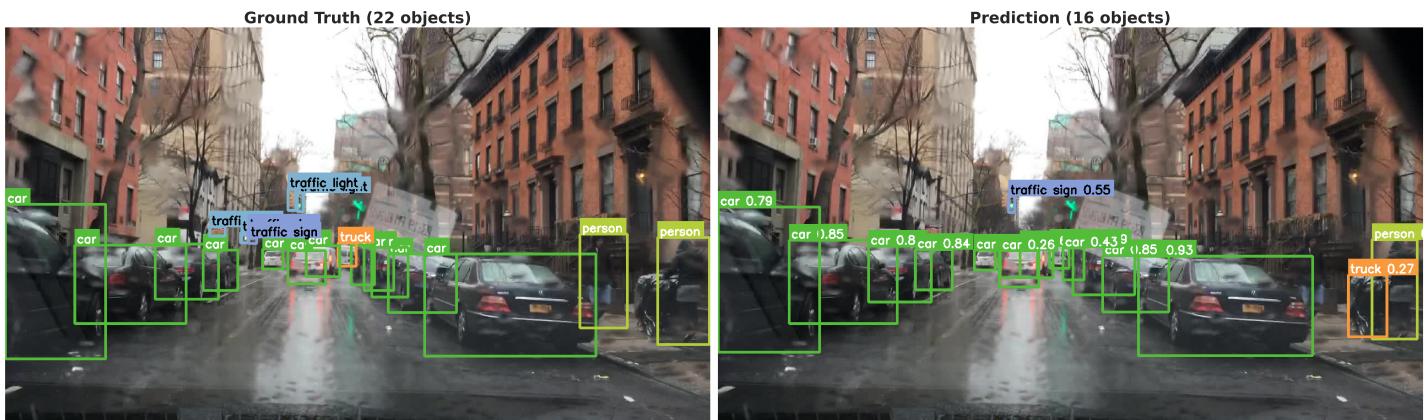
Sample #2: db0693bb-d84956a9.jpg

Ground Truth Objects: 32 | Predicted Objects: 14
Weather: clear | Scene: highway | Time of Day: daytime



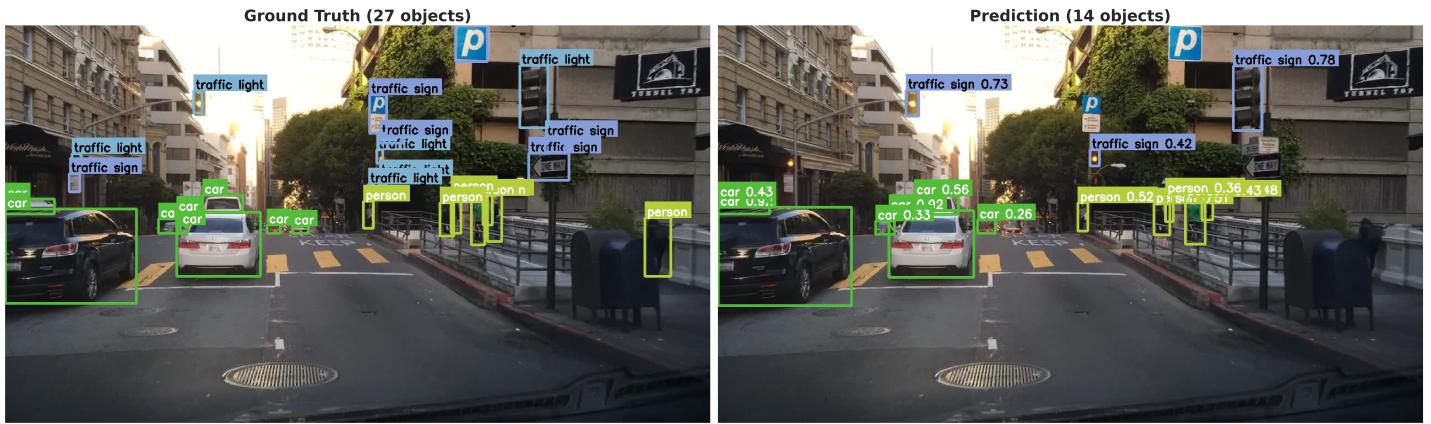
Sample #3: f4f93530-78c27314.jpg

Ground Truth Objects: 22 | Predicted Objects: 16
Weather: rainy | Scene: city street | Time of Day: daytime



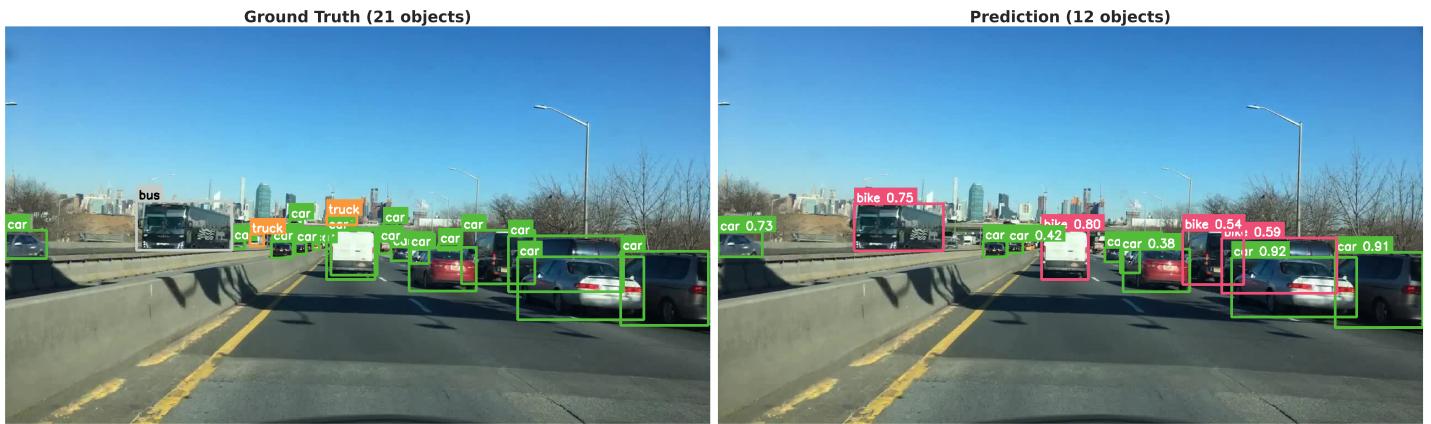
Sample #4: d87e720b-e46c3589.jpg

Ground Truth Objects: 27 | Predicted Objects: 14
Weather: undefined | Scene: city street | Time of Day: daytime



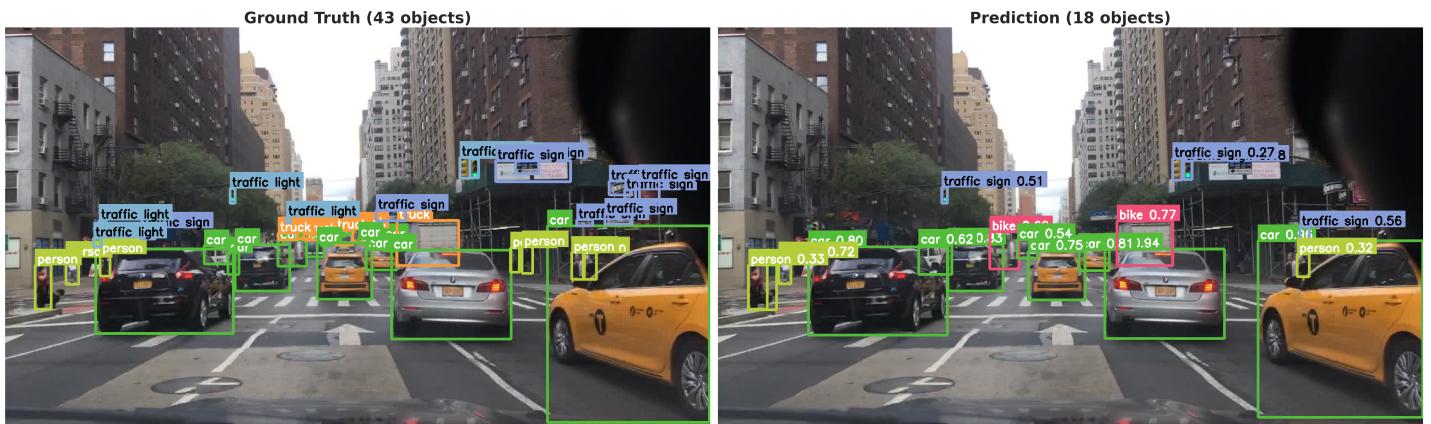
Sample #5: eebb319c-8ecffbaf.jpg

Ground Truth Objects: 21 | Predicted Objects: 12
Weather: clear | Scene: highway | Time of Day: daytime



Sample #6: d8dcae7d-b3988fc9.jpg

Ground Truth Objects: 43 | Predicted Objects: 18
Weather: undefined | Scene: city street | Time of Day: daytime



Generated by YOLO Testing Pipeline