

# YOLO Model Testing Report

<b>Model:</b>	yolo11x
<b>Model Size:</b>	109.3 MB
<b>Parameters:</b>	57.0 M
<b>FLOPs (640x640):</b>	195.96 GFLOPs
<b>Run Name:</b>	yolo11x_testing_20251128_004839
<b>Timestamp:</b>	2025-11-28 01:03:09
<b>Dataset:</b>	bdd100k_yolo_limited - test split
<b>Images Processed:</b>	20000
<b>Total Execution Time</b>	152.03s
<b>IoU Threshold:</b>	0.5

## Inference Performance

Metric	Value
Average Inference Time	4.21 ms per image
FPS (Frames Per Second)	237.60

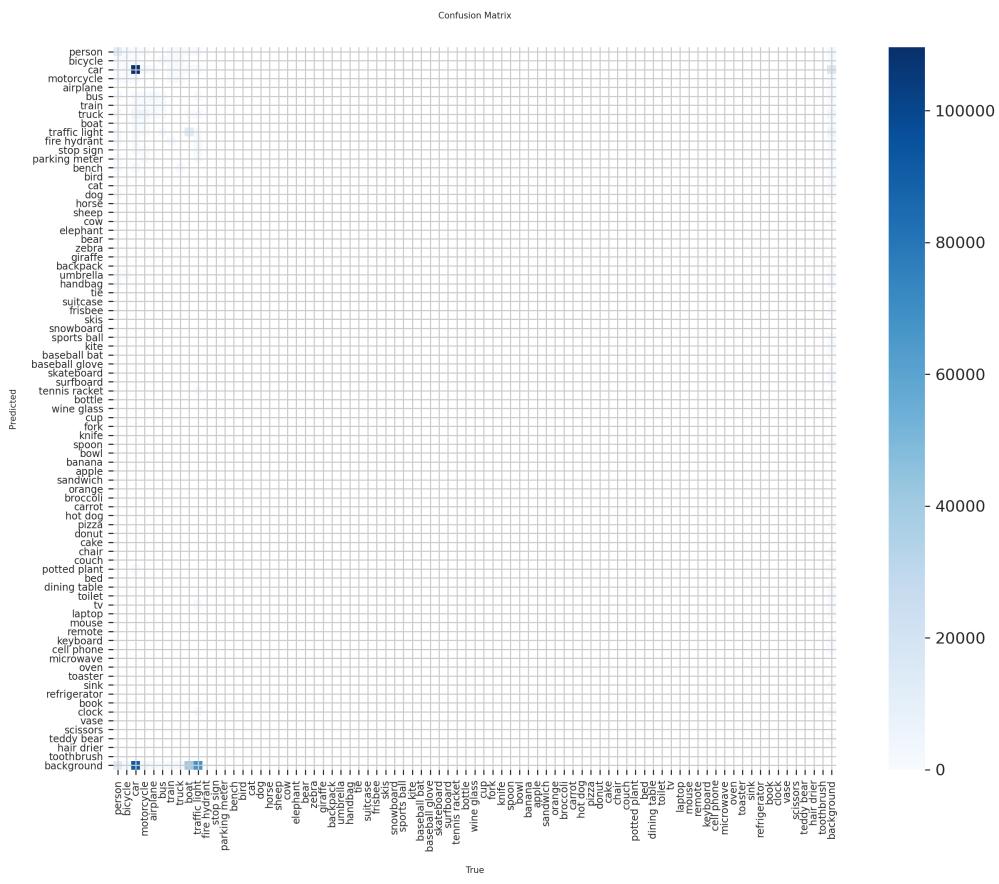
Note: Inference time includes preprocessing, model inference, and postprocessing.

## Overall Accuracy Metrics

Metric	Value
Precision	0.3295
Recall	0.7026
F1-Score	0.4486
mAP@0.5	0.1310
mAP@0.5:0.95	0.0730
True Positives	121153
False Positives	246574
False Negatives	51279

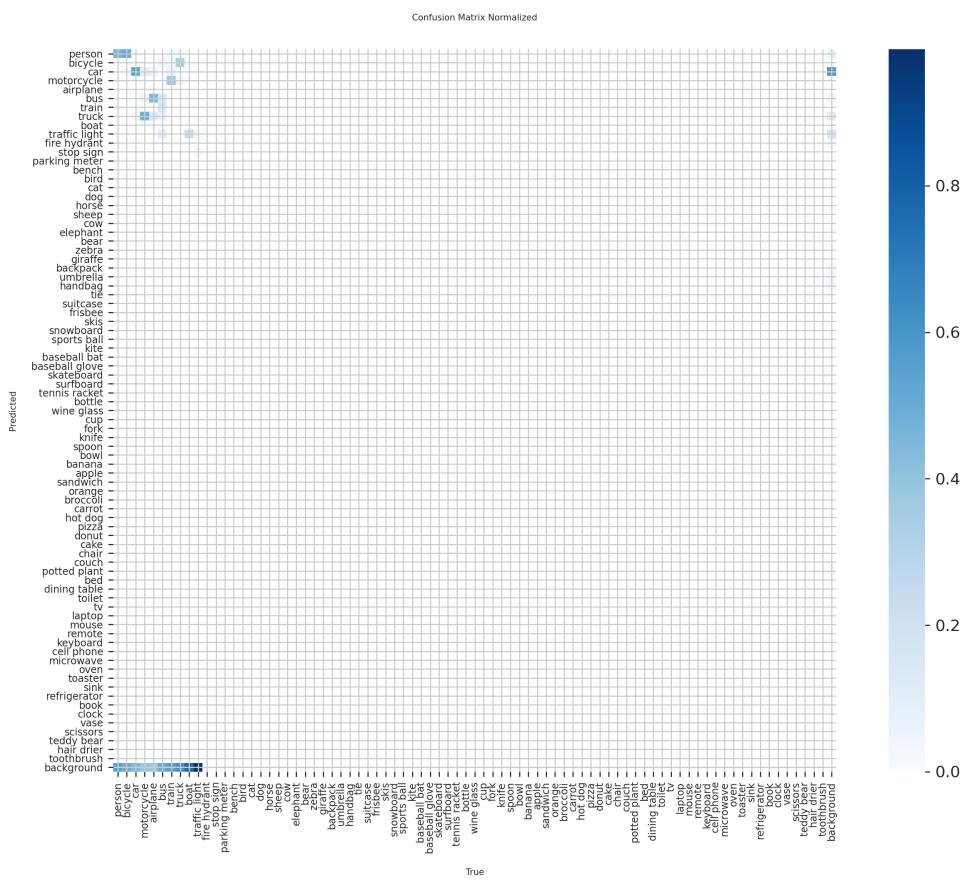
# Confusion Matrices

## Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

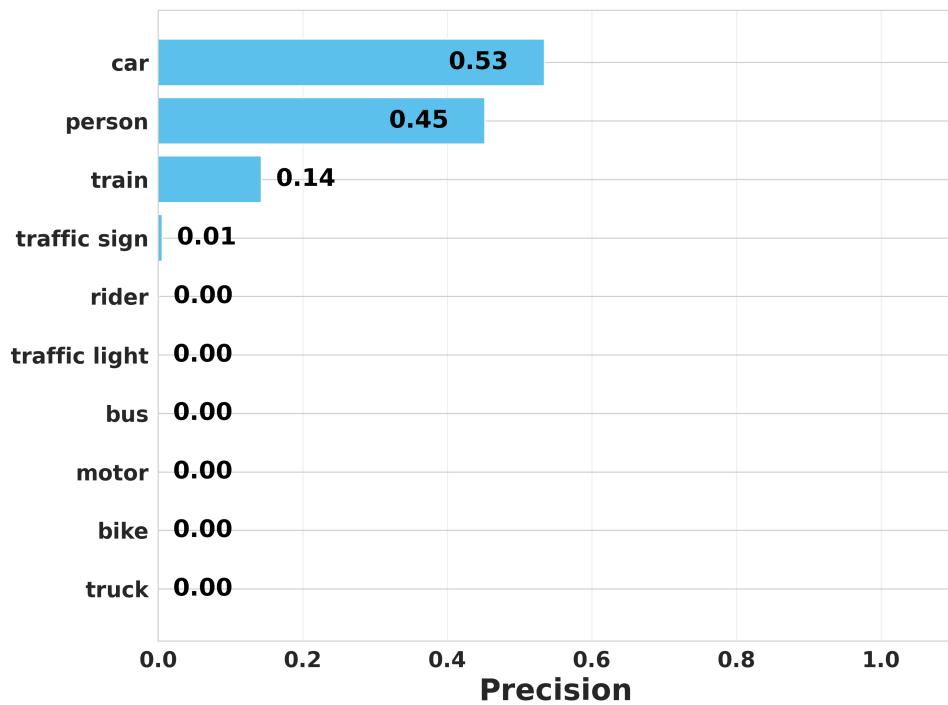
## Confusion Matrix (Normalized)



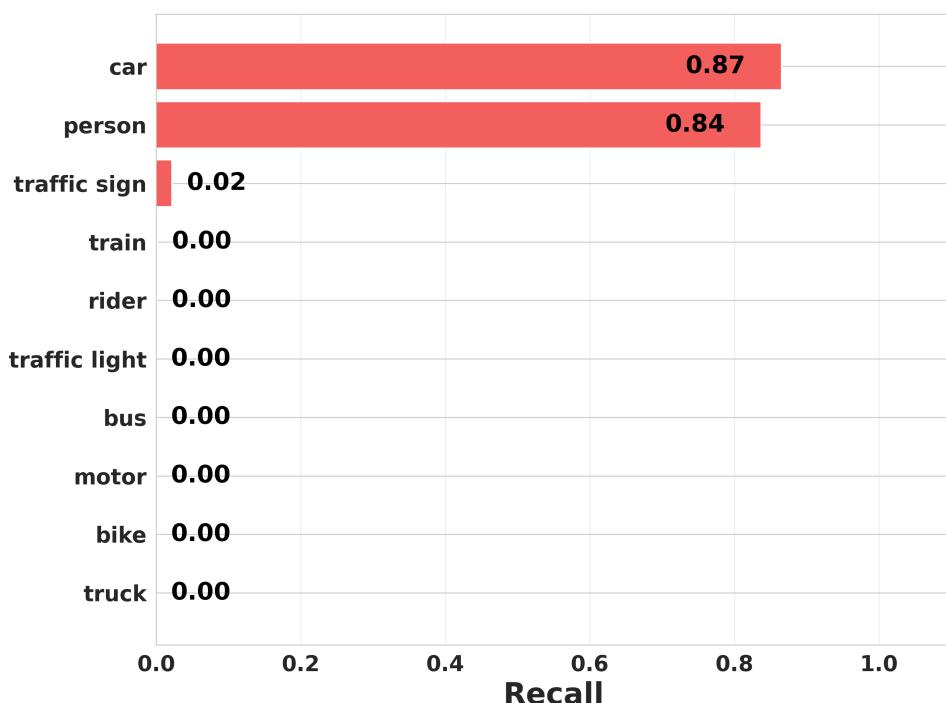
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

# Performance Visualizations

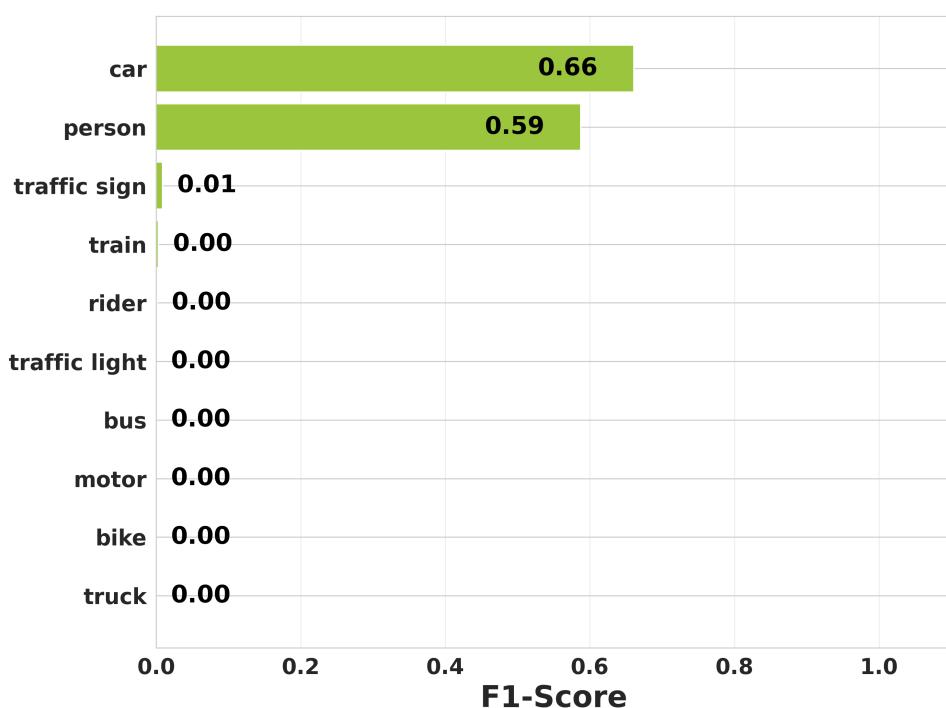
## Precision by Class



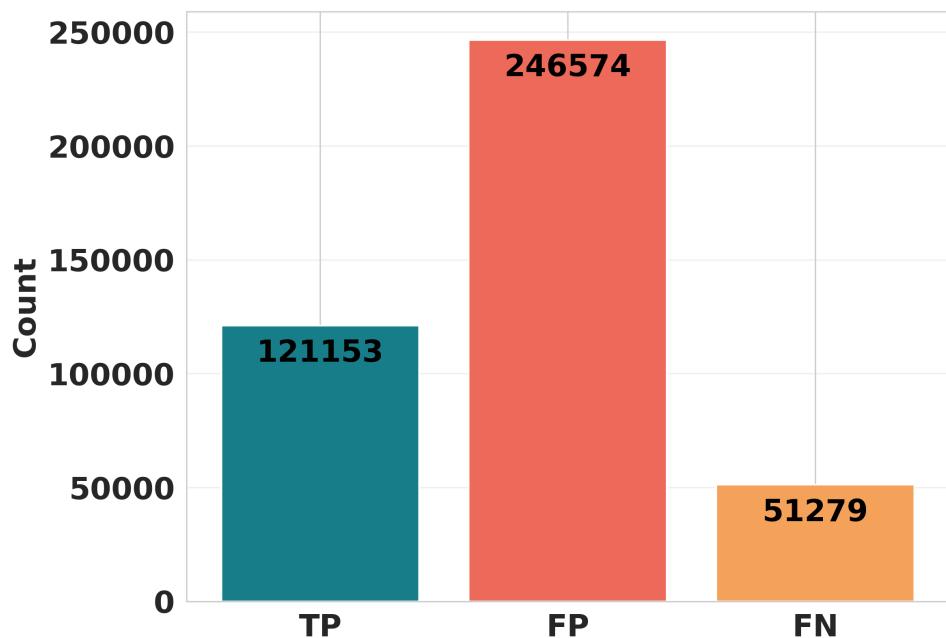
### Recall by Class



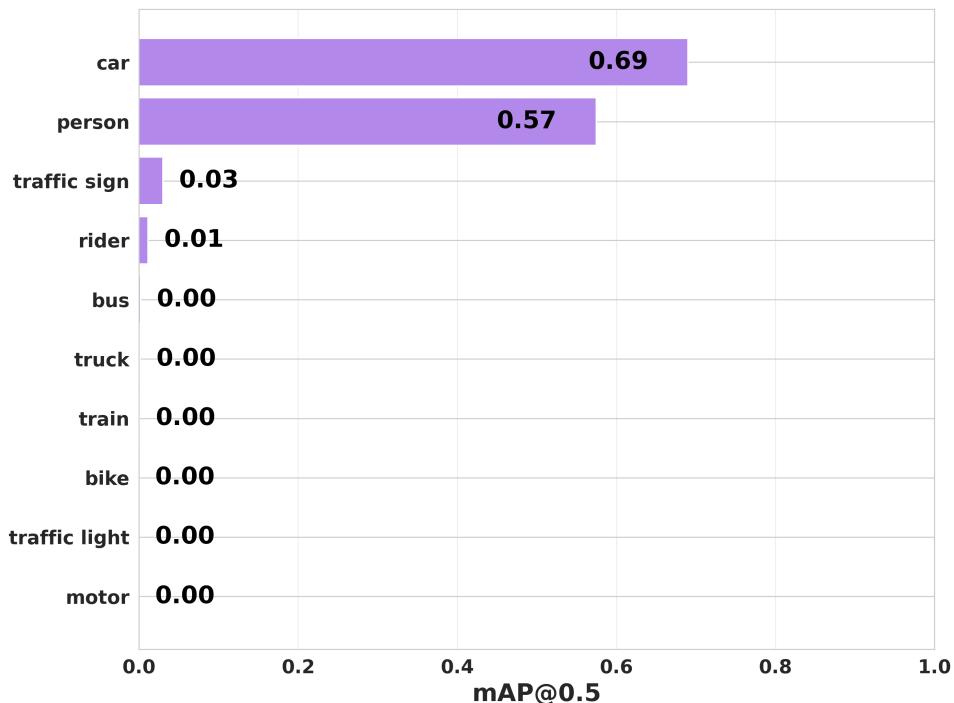
### F1-Score by Class



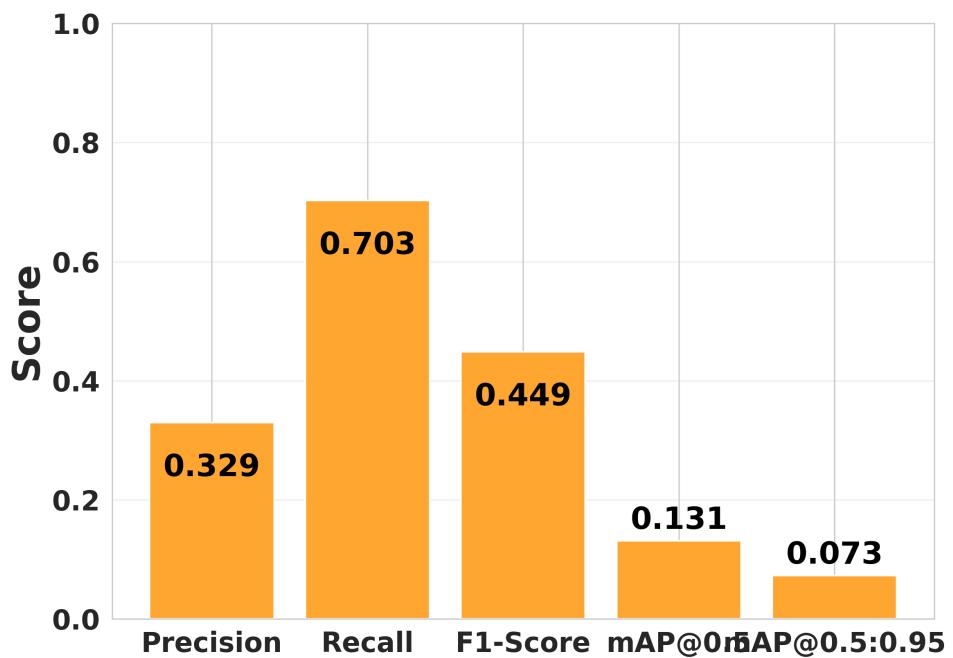
## Overall Detection Outcomes



## mAP@0.5 by Class



## Overall Metrics



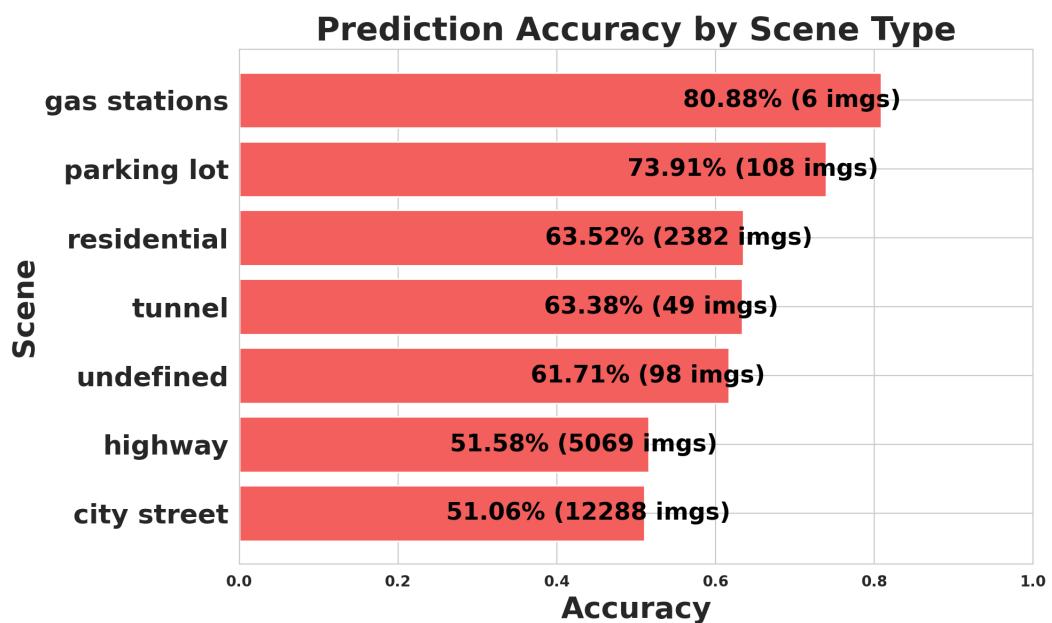
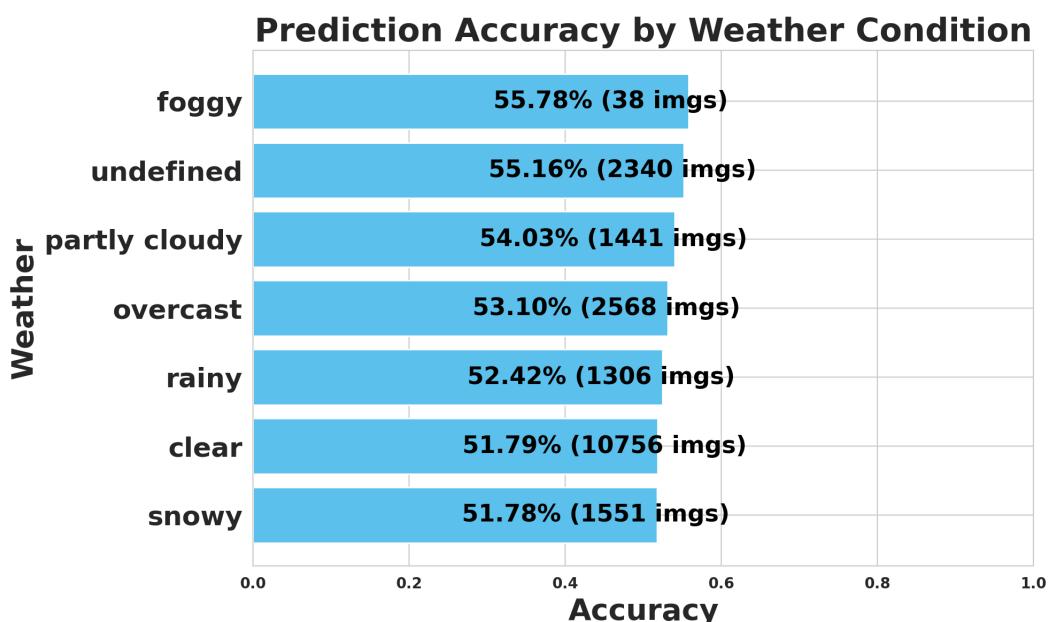
## Per-Class Performance

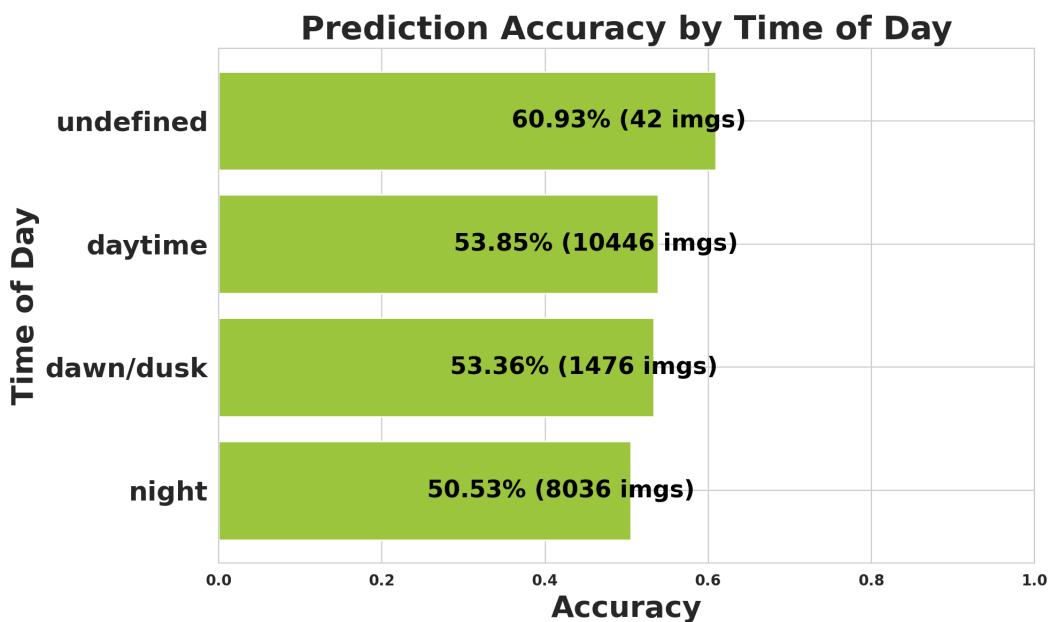
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	11145	13505	2168	0.4521	0.8372	0.5872	0.5748
rider	1	1293	815	0.0008	0.0012	0.0009	0.0112
car	109626	95523	17099	0.5344	0.8651	0.6606	0.6900
truck	0	8704	462	0.0000	0.0000	0.0000	0.0015
bus	0	3217	15	0.0000	0.0000	0.0000	0.0020
train	4	24	2811	0.1429	0.0014	0.0028	0.0006
motor	0	841	120	0.0000	0.0000	0.0000	0.0000
bike	0	1998	10946	0.0000	0.0000	0.0000	0.0001
traffic light	0	52840	15	0.0000	0.0000	0.0000	0.0000
traffic sign	377	68629	16828	0.0055	0.0219	0.0087	0.0300

# Comprehensive Failure Analysis

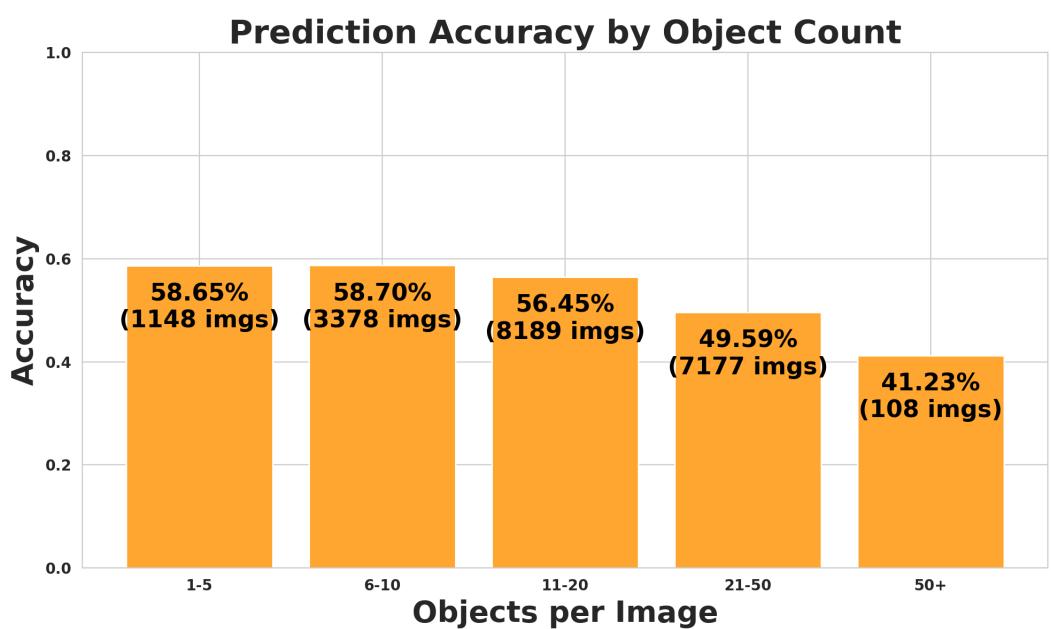
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	193533
Overall Accuracy	52.63%

## Accuracy by Environmental Attributes

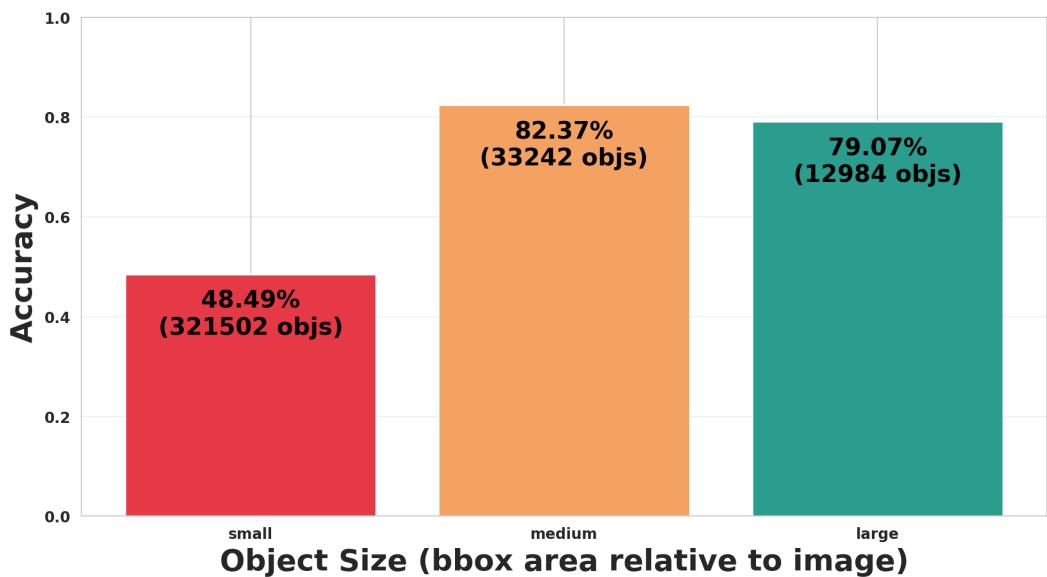




## Accuracy by Object Characteristics

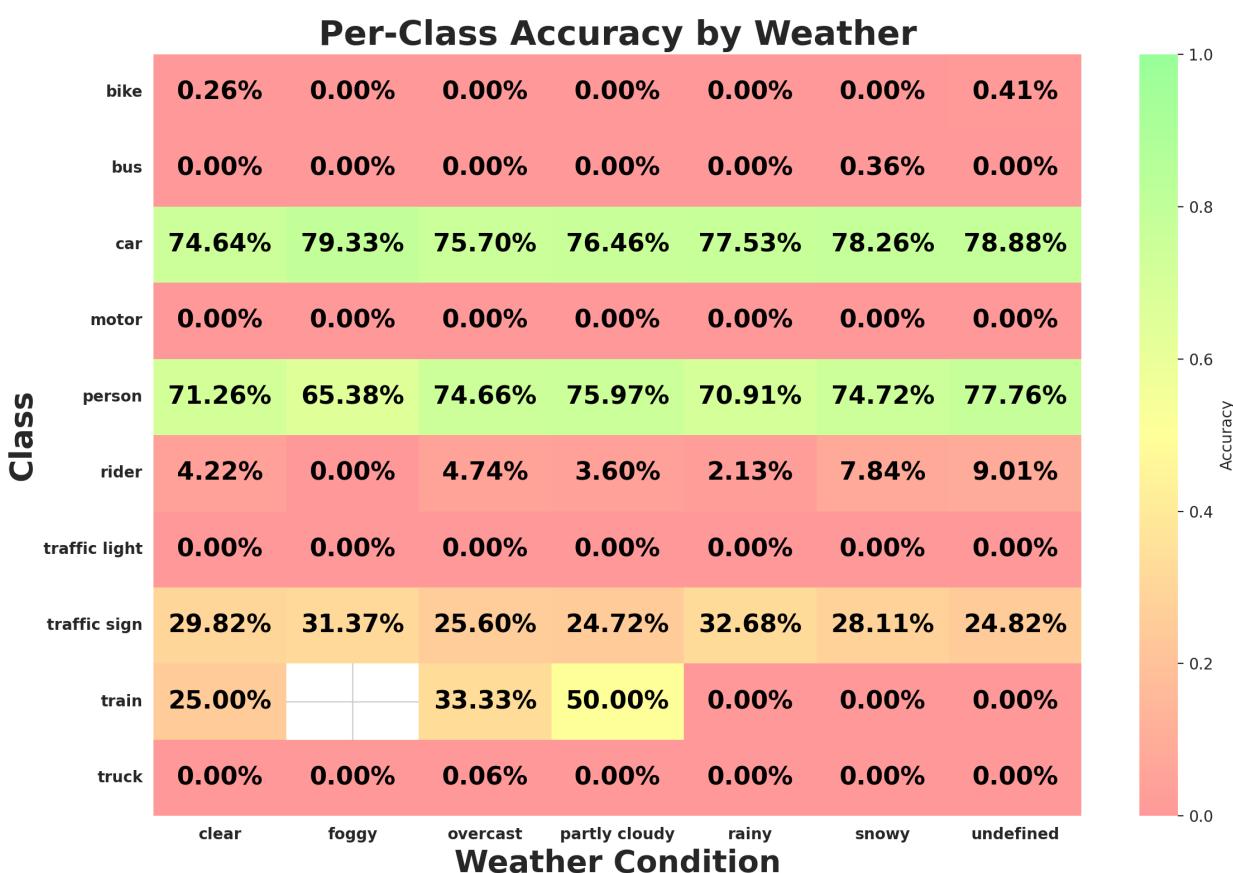
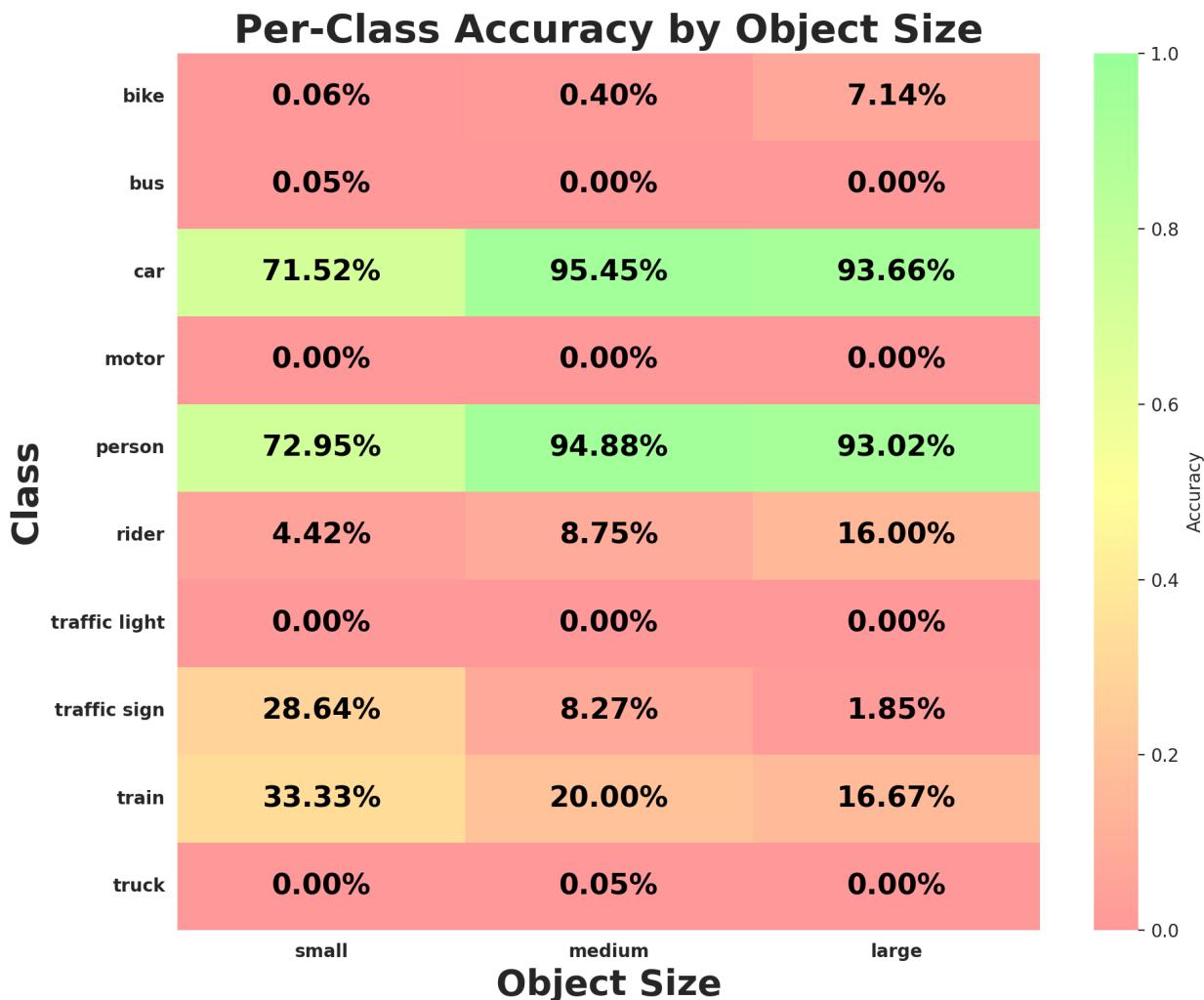


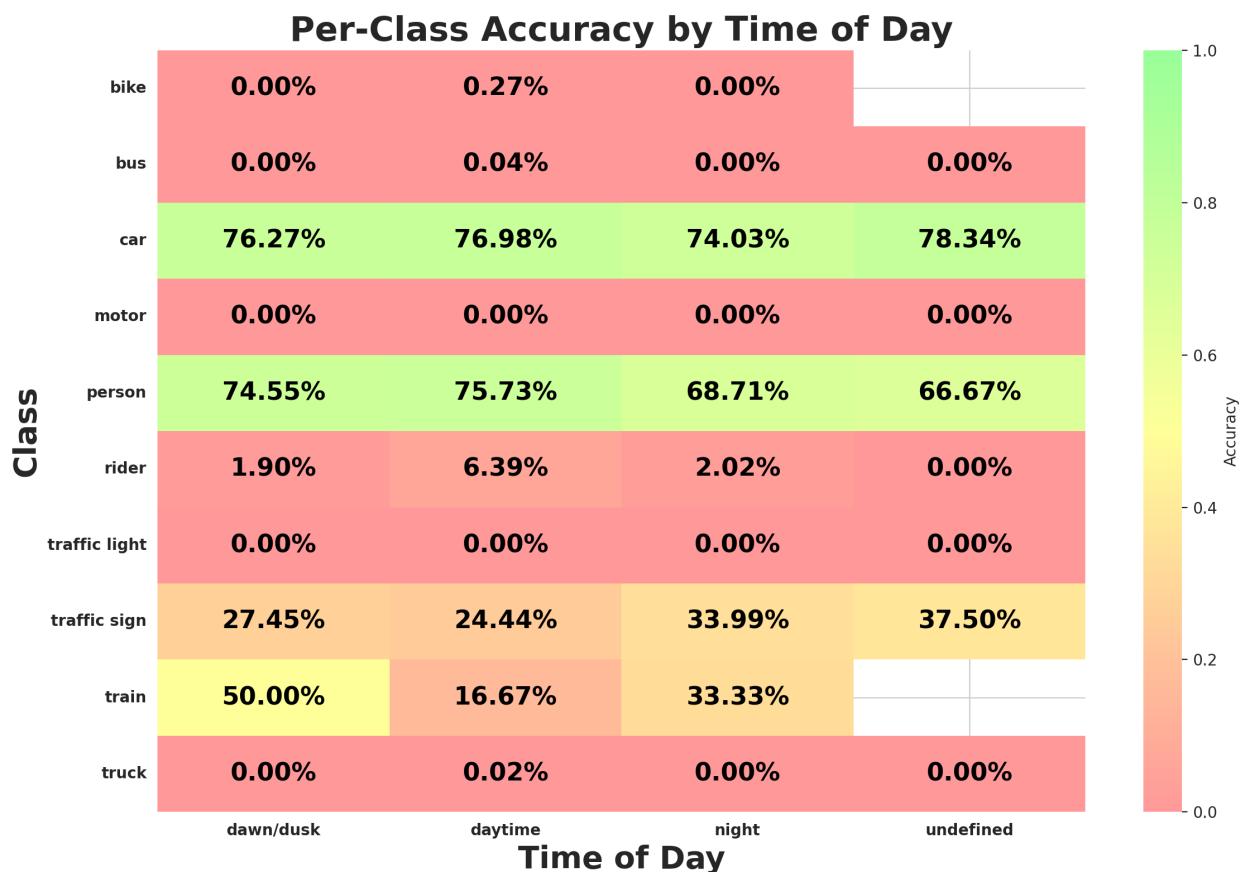
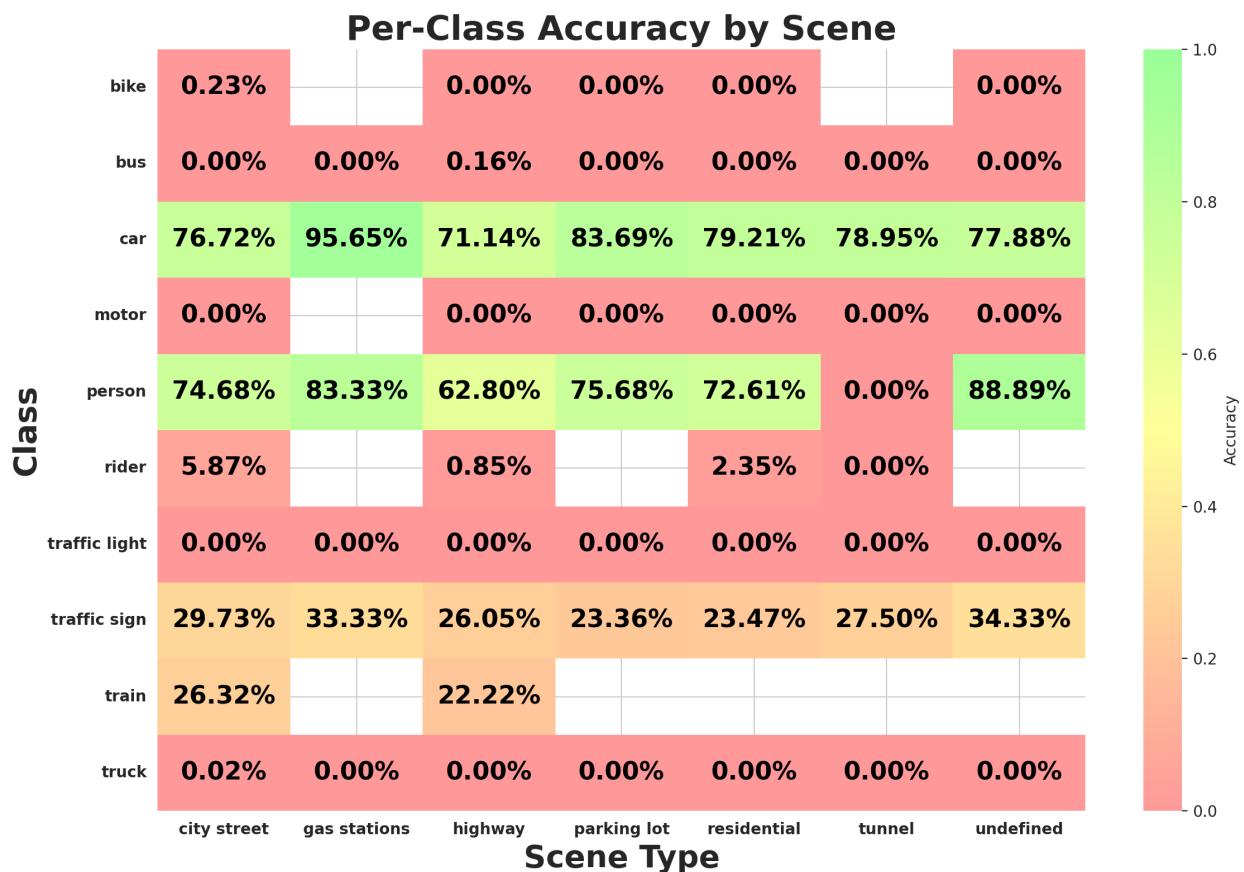
### Prediction Accuracy by Object Size



**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



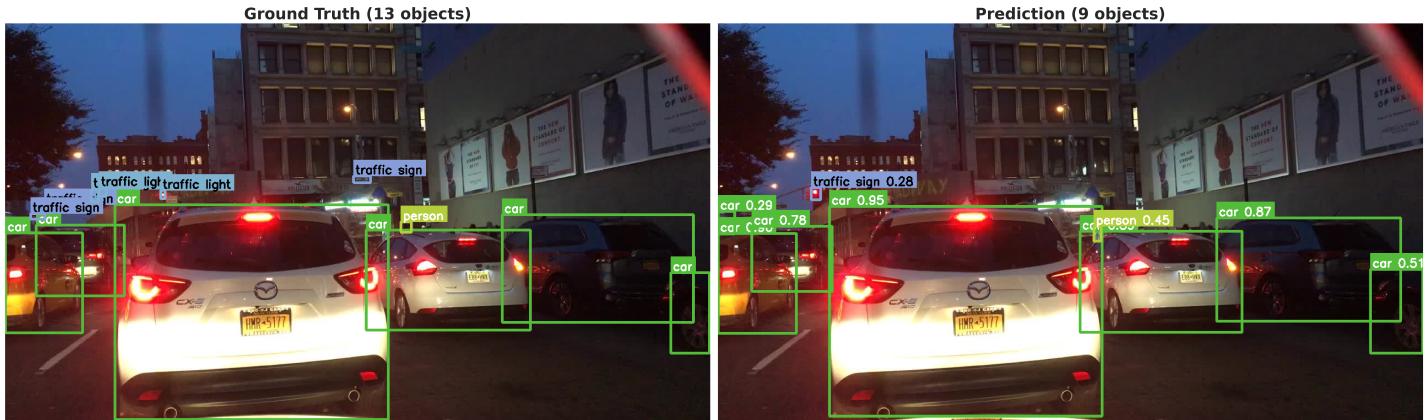


# Sample Predictions: Ground Truth vs Model

## Sample #1: fd2f2d25-04cca5b4.jpg

Ground Truth Objects: 13 | Predicted Objects: 9

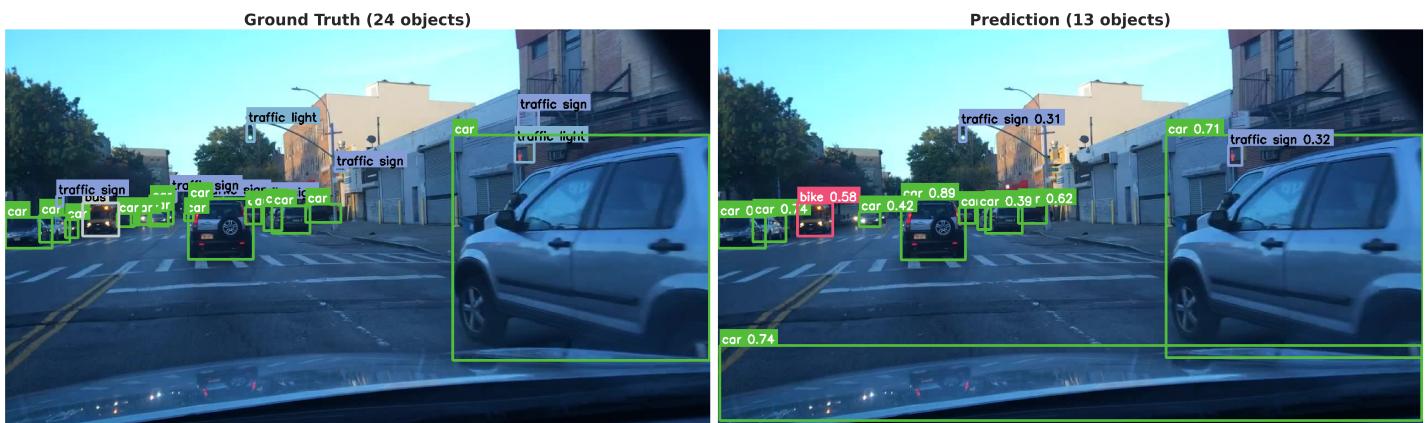
Weather: clear | Scene: city street | Time of Day: dawn/dusk



## Sample #2: f8a48562-dd428ddc.jpg

Ground Truth Objects: 24 | Predicted Objects: 13

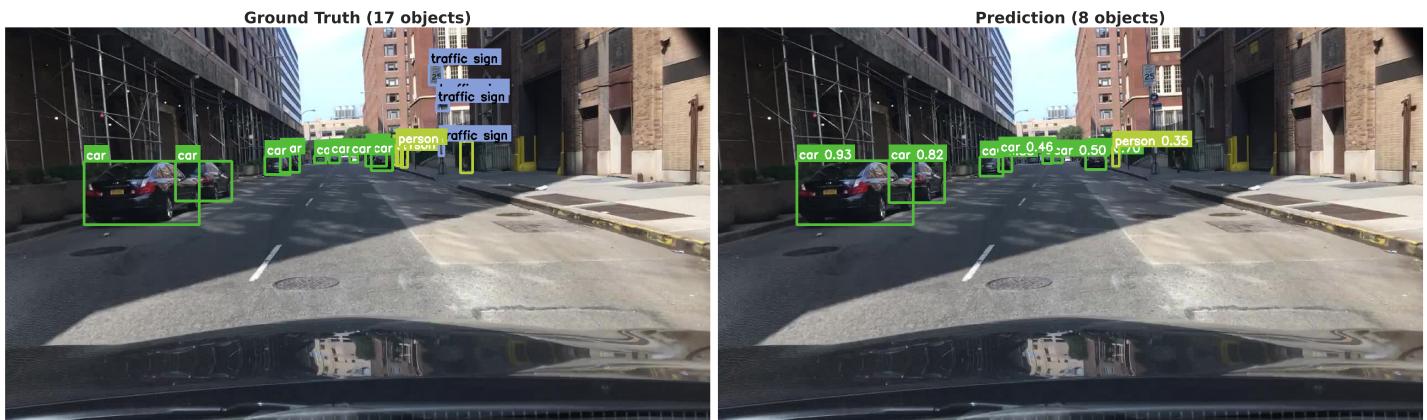
Weather: clear | Scene: city street | Time of Day: daytime



## Sample #3: cc55cc0b-e3bd032c.jpg

Ground Truth Objects: 17 | Predicted Objects: 8

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



## Sample #4: f7e0f4ad-8b619570.jpg

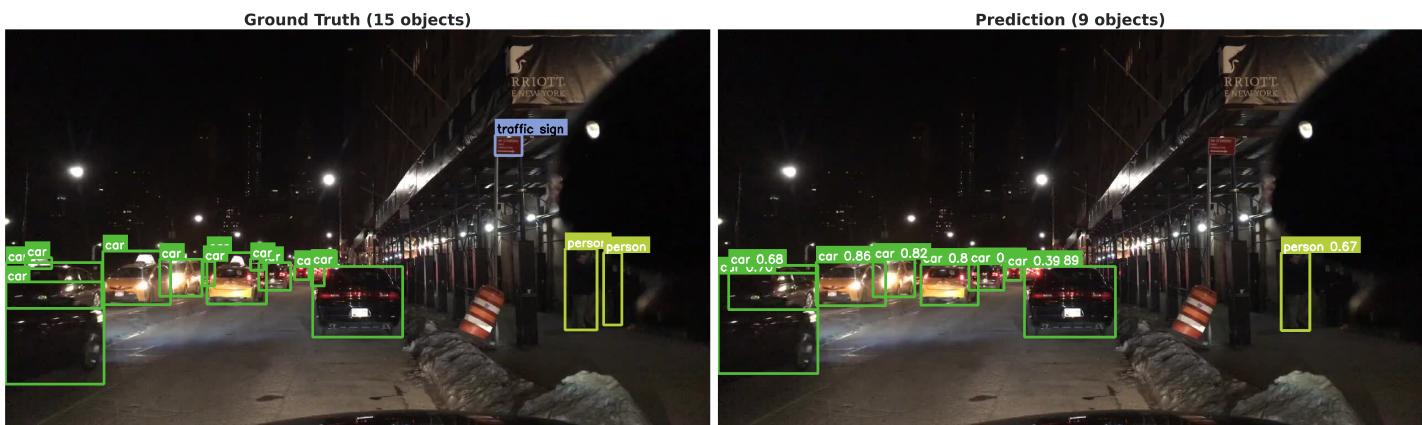
Ground Truth Objects: 10 | Predicted Objects: 6

Weather: clear | Scene: highway | Time of Day: night



**Sample #5: d654301b-5583bb1d.jpg**

Ground Truth Objects: 15 | Predicted Objects: 9  
Weather: snowy | Scene: city street | Time of Day: night



**Sample #6: d49af90b-061c1a95.jpg**

Ground Truth Objects: 16 | Predicted Objects: 10  
Weather: partly cloudy | Scene: city street | Time of Day: daytime



Generated by YOLO Testing Pipeline