

# YOLO Model Testing Report

<b>Model:</b>	yolo11s
<b>Model Size:</b>	18.4 MB
<b>Parameters:</b>	9.5 M
<b>FLOPs (640x640):</b>	21.72 GFLOPs
<b>Run Name:</b>	yolo11s_testing_20251128_002304
<b>Timestamp:</b>	2025-11-28 00:35:15
<b>Dataset:</b>	bdd100k_yolo_limited - test split
<b>Images Processed:</b>	20000
<b>Total Execution Time</b>	105.41s
<b>IoU Threshold:</b>	0.5

## Inference Performance

Metric	Value
Average Inference Time	1.81 ms per image
FPS (Frames Per Second)	552.49

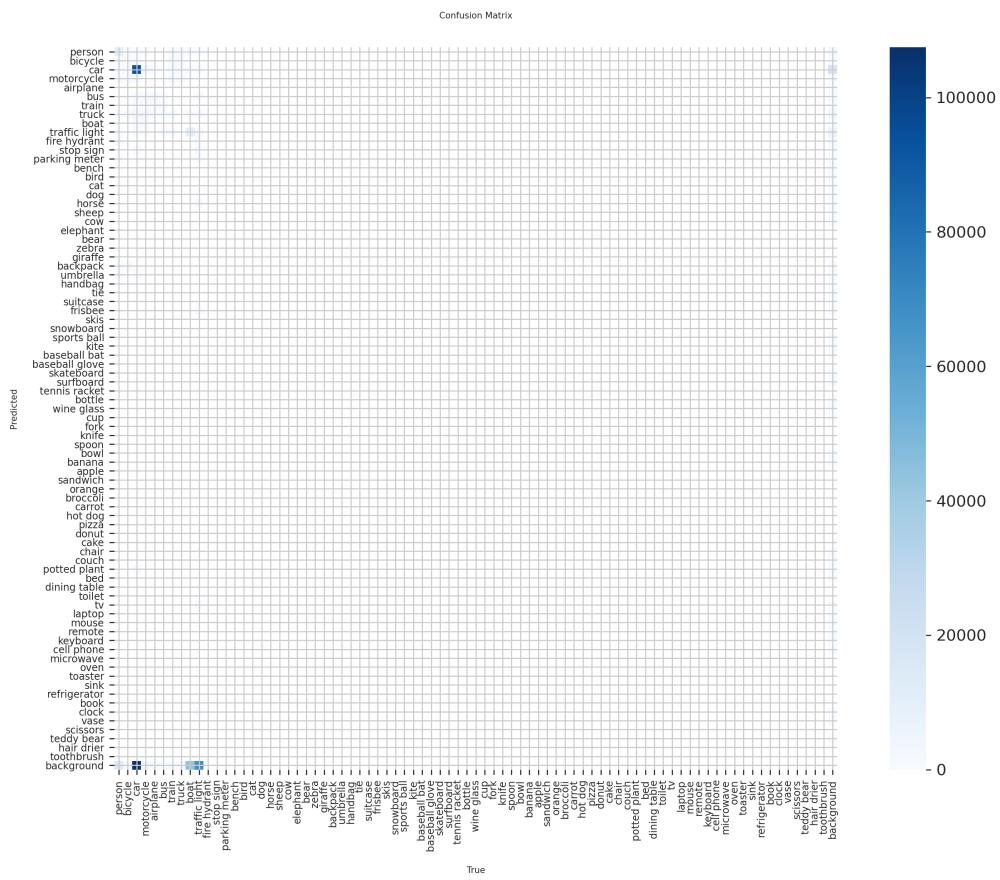
Note: Inference time includes preprocessing, model inference, and postprocessing.

## Overall Accuracy Metrics

Metric	Value
Precision	0.2843
Recall	0.6712
F1-Score	0.3994
mAP@0.5	0.1119
mAP@0.5:0.95	0.0598
True Positives	104535
False Positives	263192
False Negatives	51219

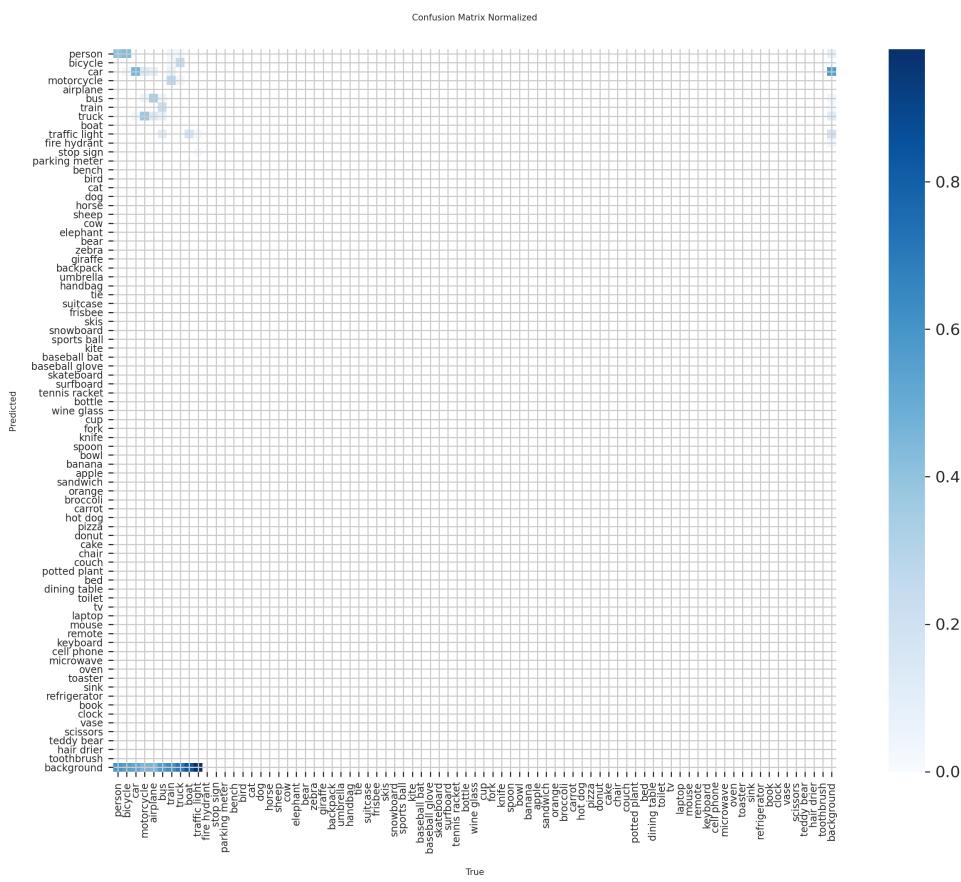
# Confusion Matrices

## Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

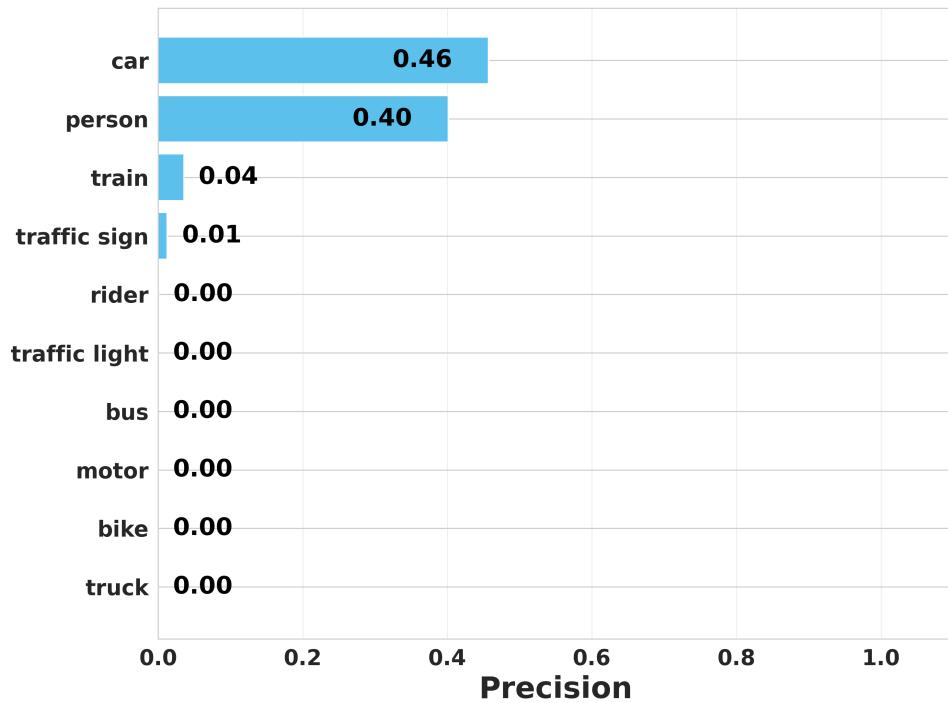
## Confusion Matrix (Normalized)



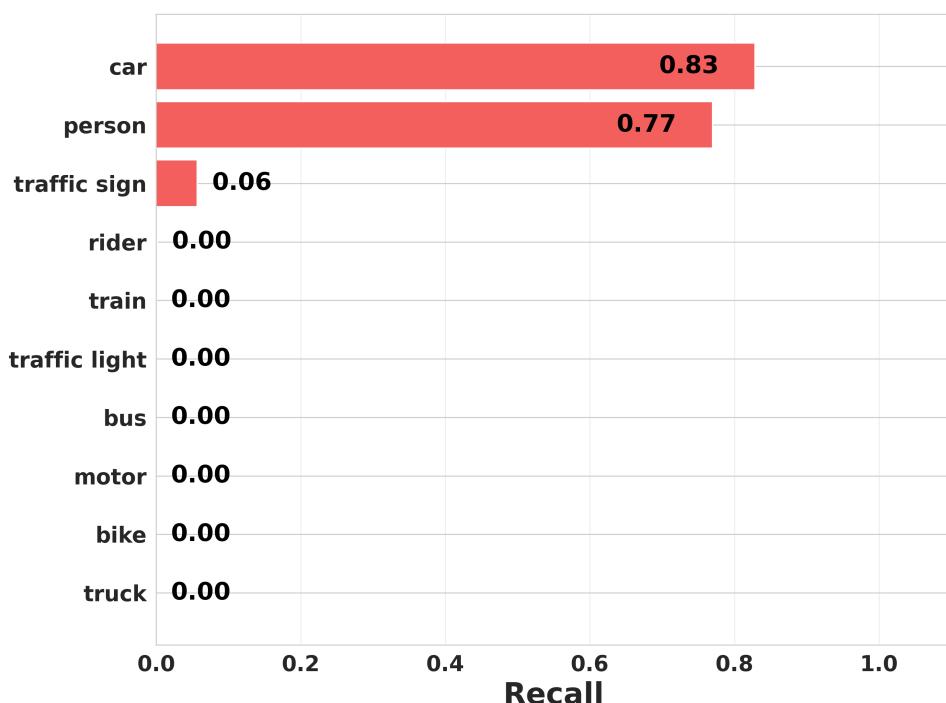
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

# Performance Visualizations

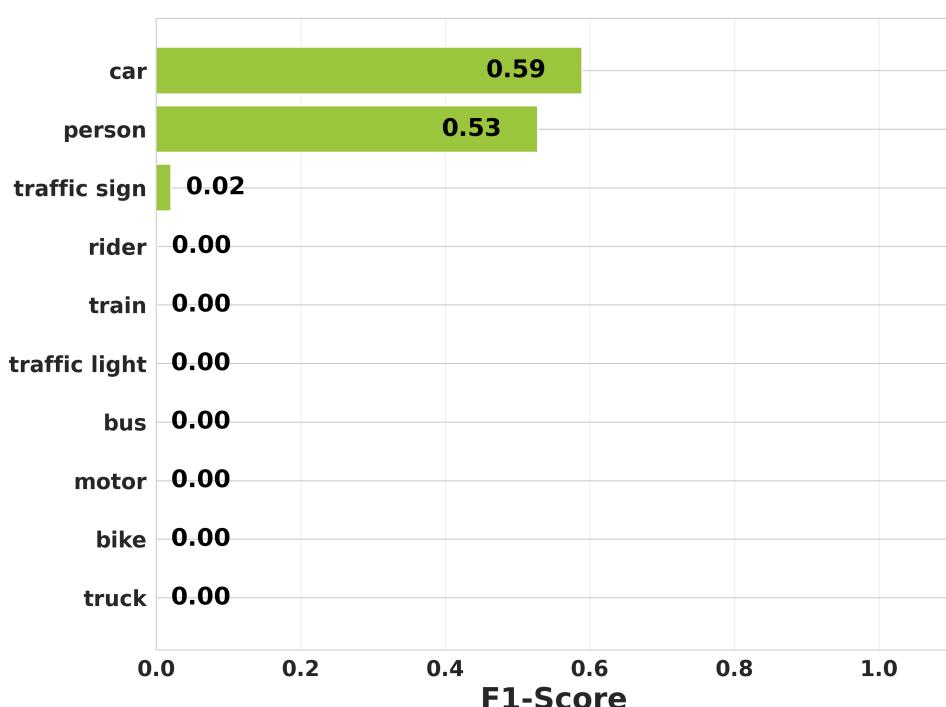
## Precision by Class



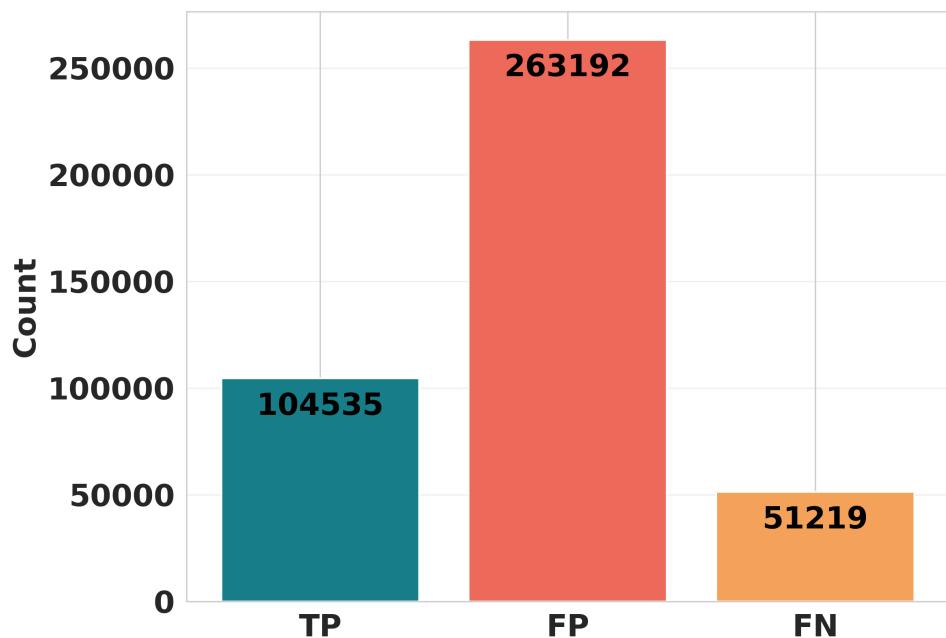
### Recall by Class



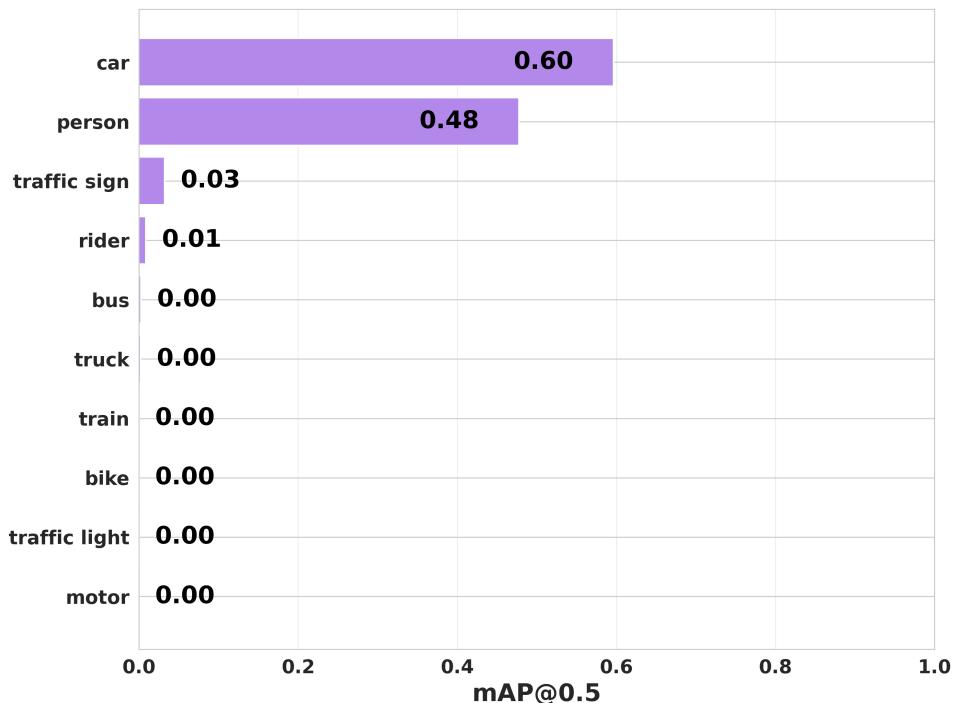
### F1-Score by Class



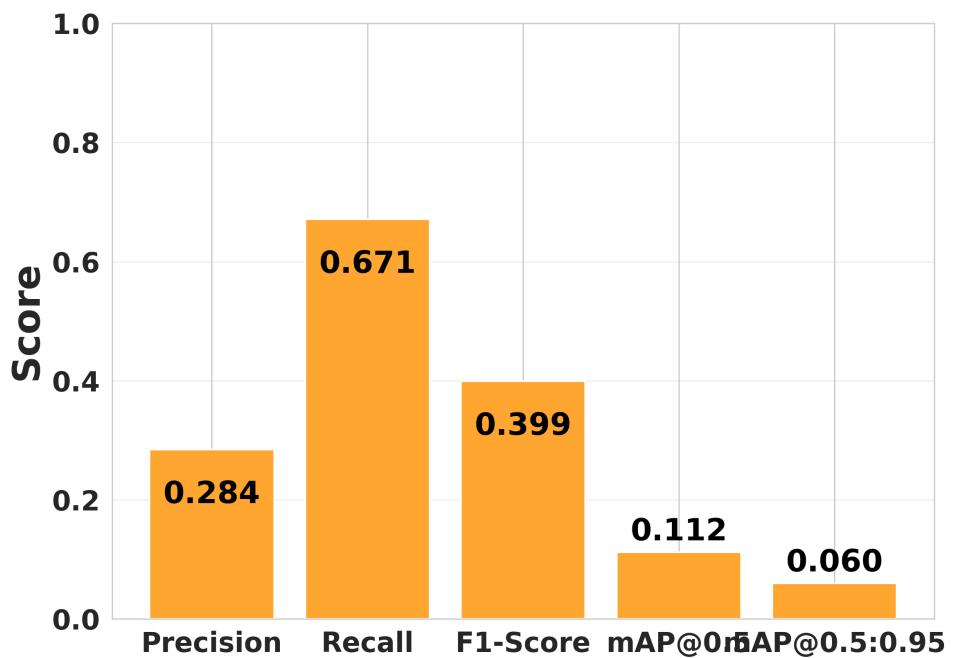
## Overall Detection Outcomes



## mAP@0.5 by Class



## Overall Metrics



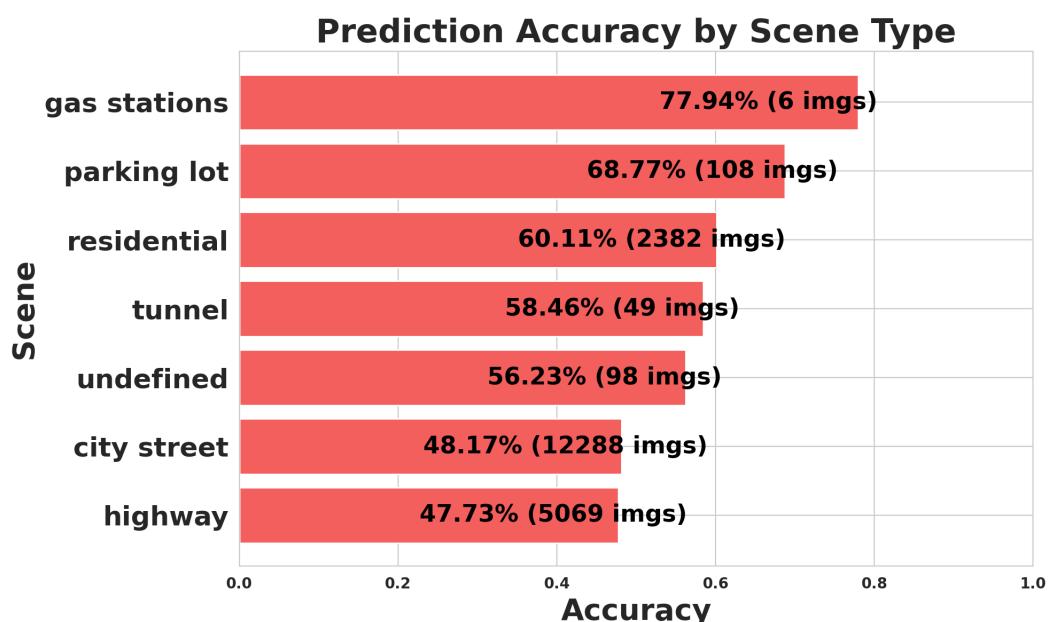
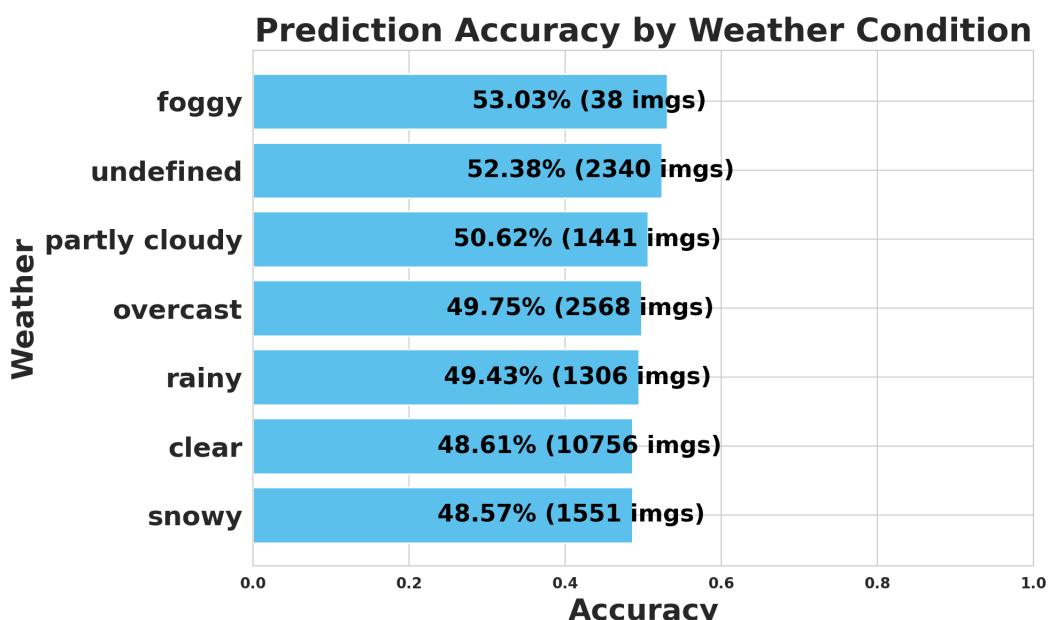
## Per-Class Performance

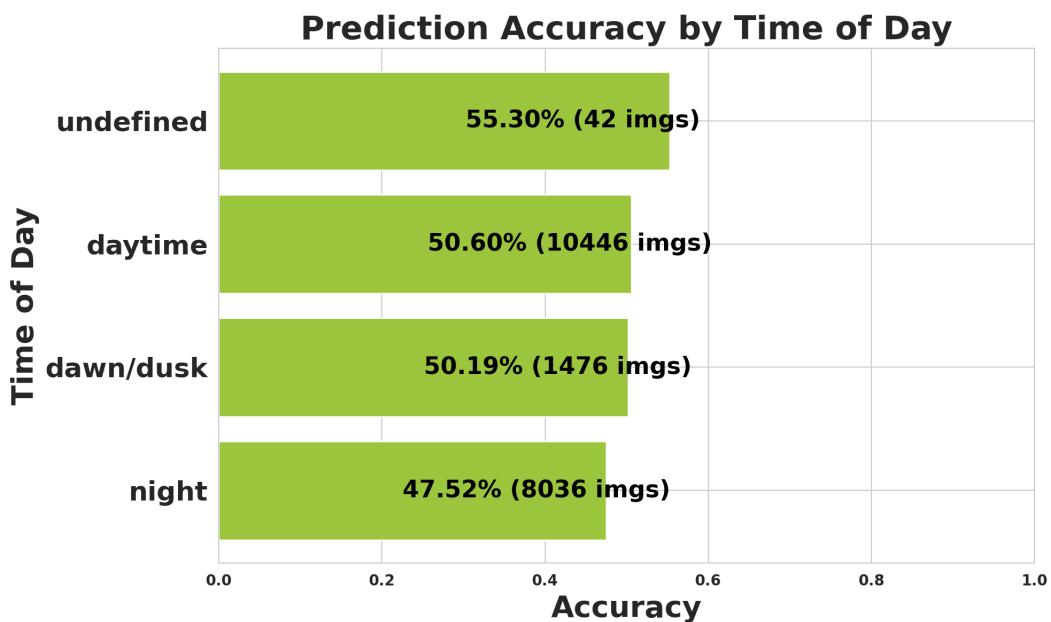
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	9896	14754	2957	0.4015	0.7699	0.5277	0.4774
rider	1	1293	669	0.0008	0.0015	0.0010	0.0083
car	93772	111377	19426	0.4571	0.8284	0.5891	0.5962
truck	0	8704	456	0.0000	0.0000	0.0000	0.0019
bus	0	3217	20	0.0000	0.0000	0.0000	0.0023
train	1	27	2733	0.0357	0.0004	0.0007	0.0002
motor	0	841	181	0.0000	0.0000	0.0000	0.0000
bike	0	1998	10418	0.0000	0.0000	0.0000	0.0001
traffic light	0	52840	13	0.0000	0.0000	0.0000	0.0001
traffic sign	865	68141	14346	0.0125	0.0569	0.0205	0.0320

# Comprehensive Failure Analysis

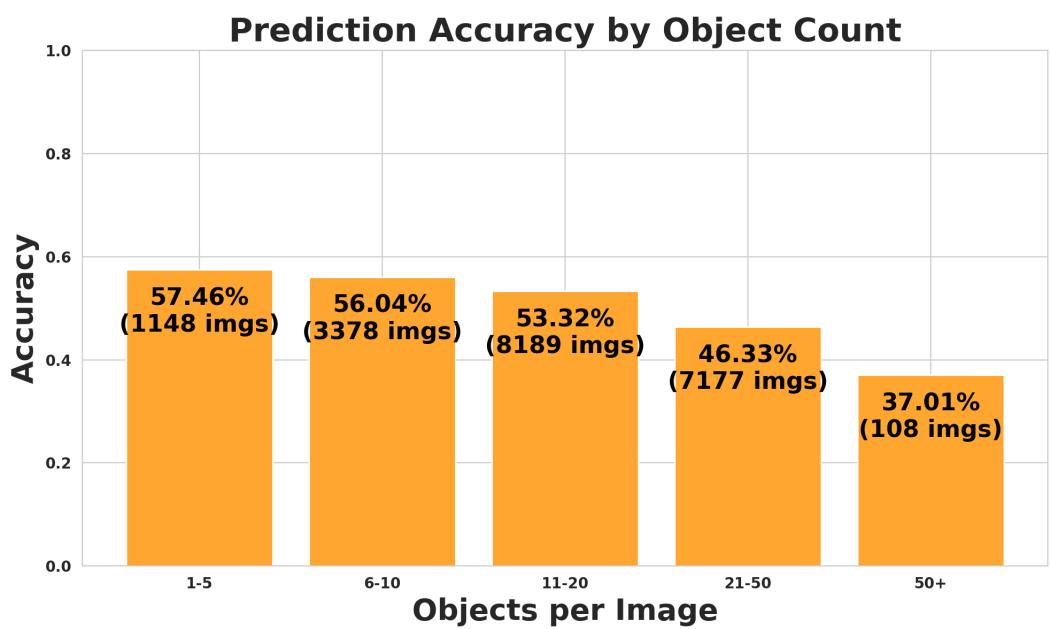
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	181913
Overall Accuracy	49.47%

## Accuracy by Environmental Attributes

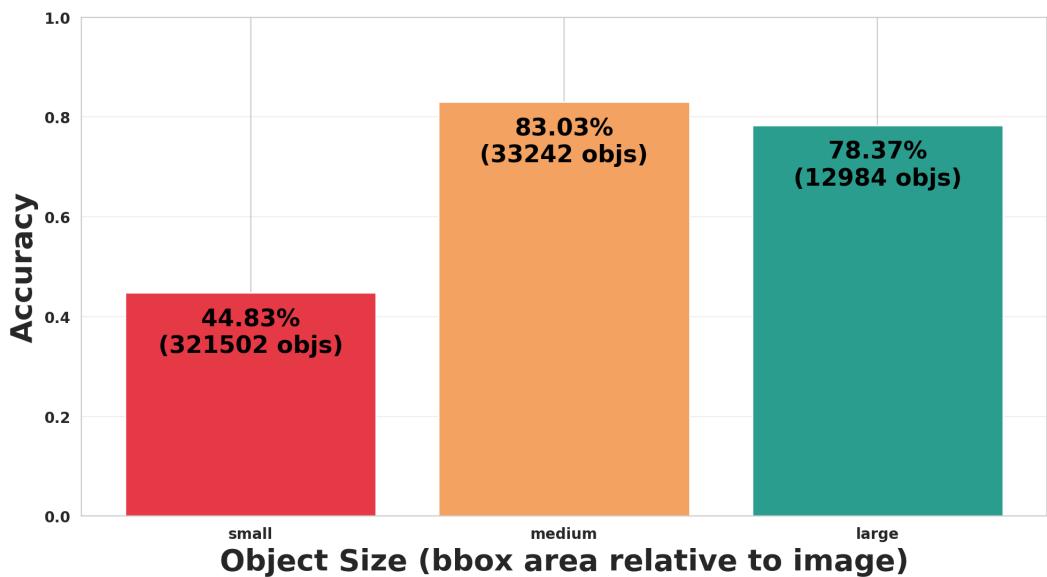




## Accuracy by Object Characteristics

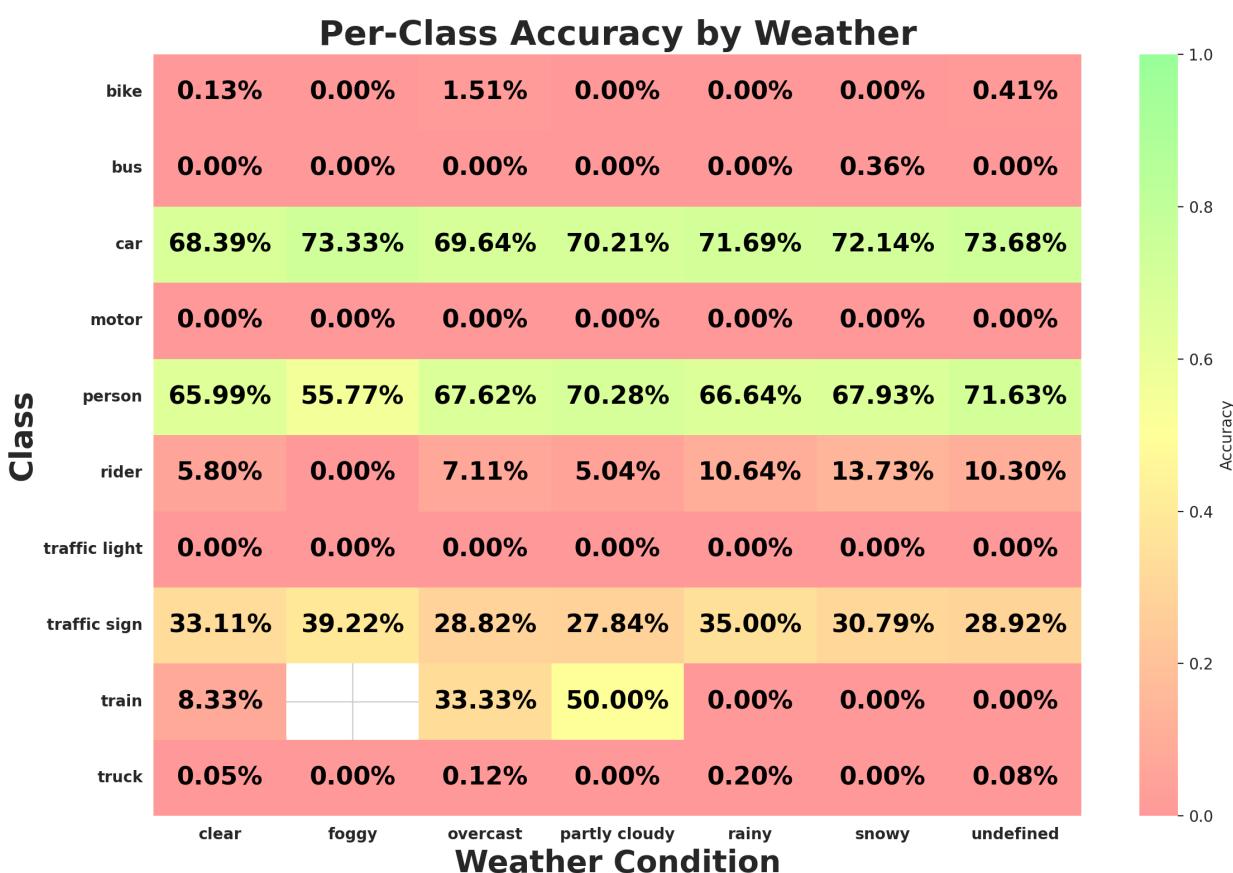
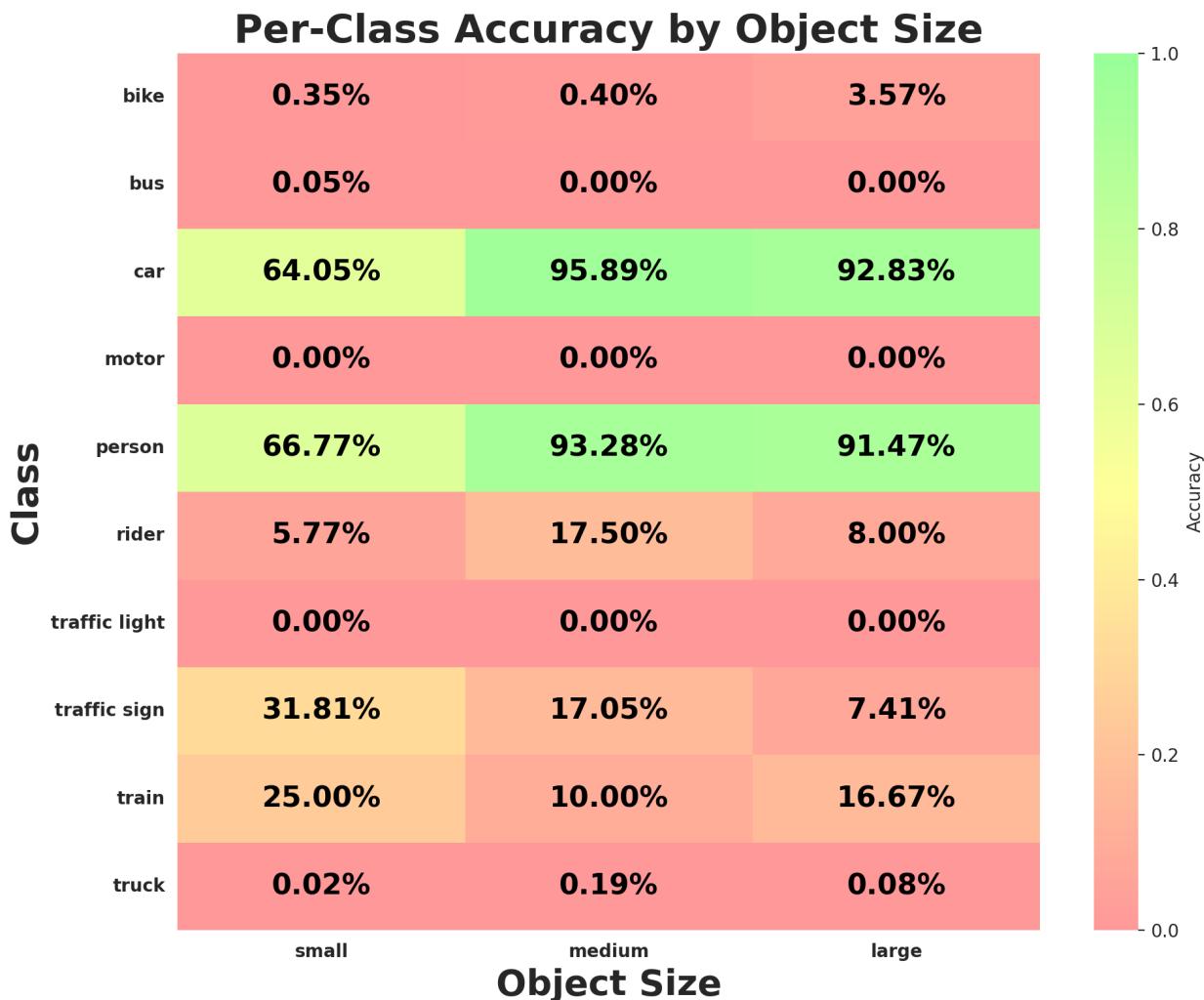


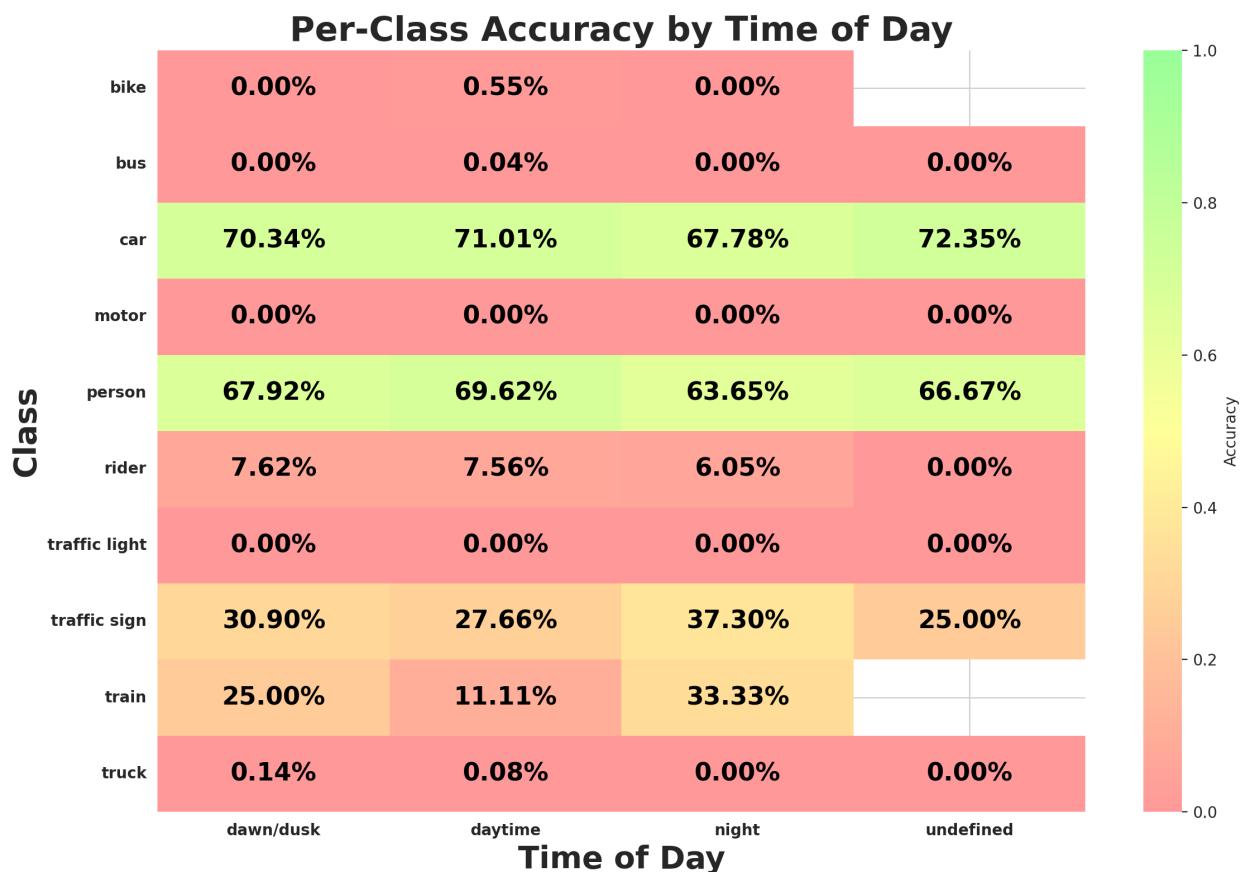
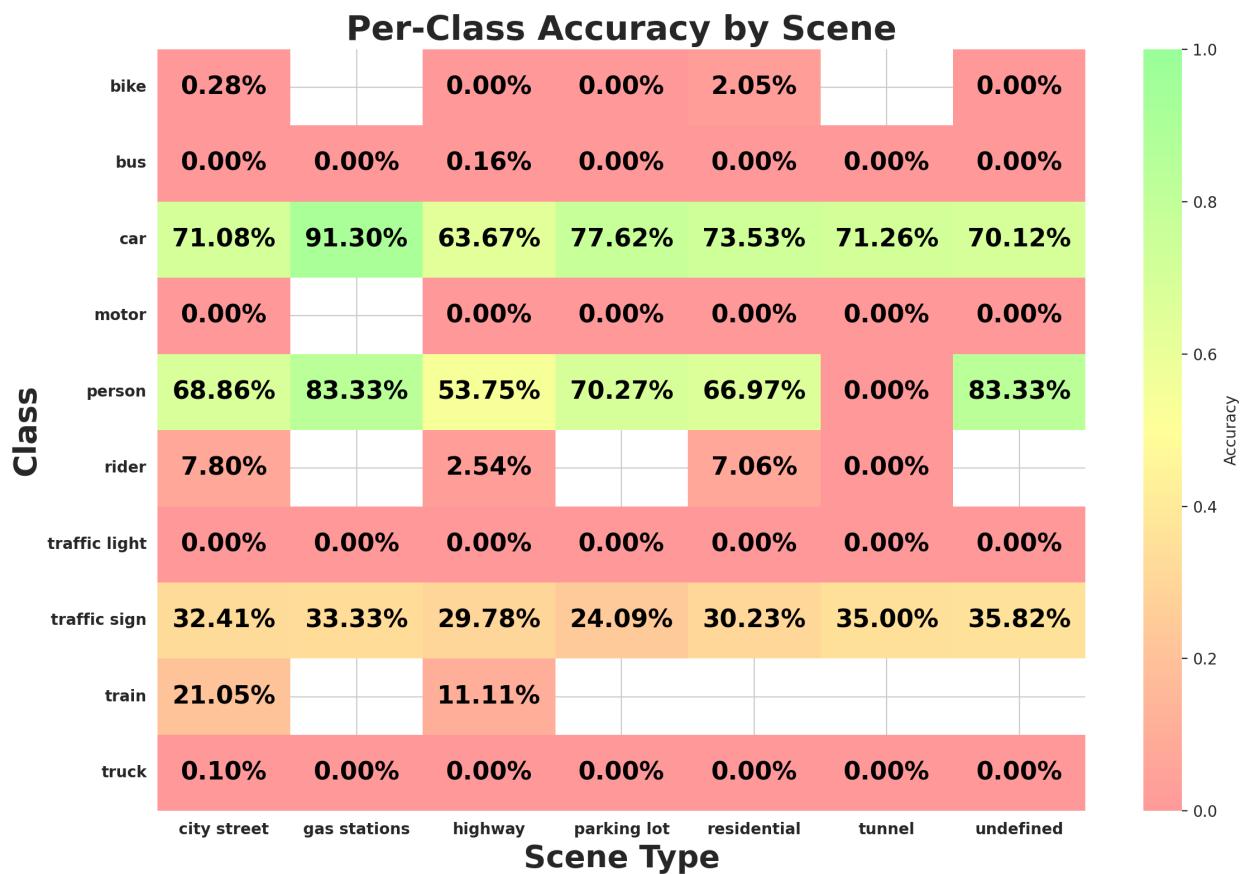
### Prediction Accuracy by Object Size



**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

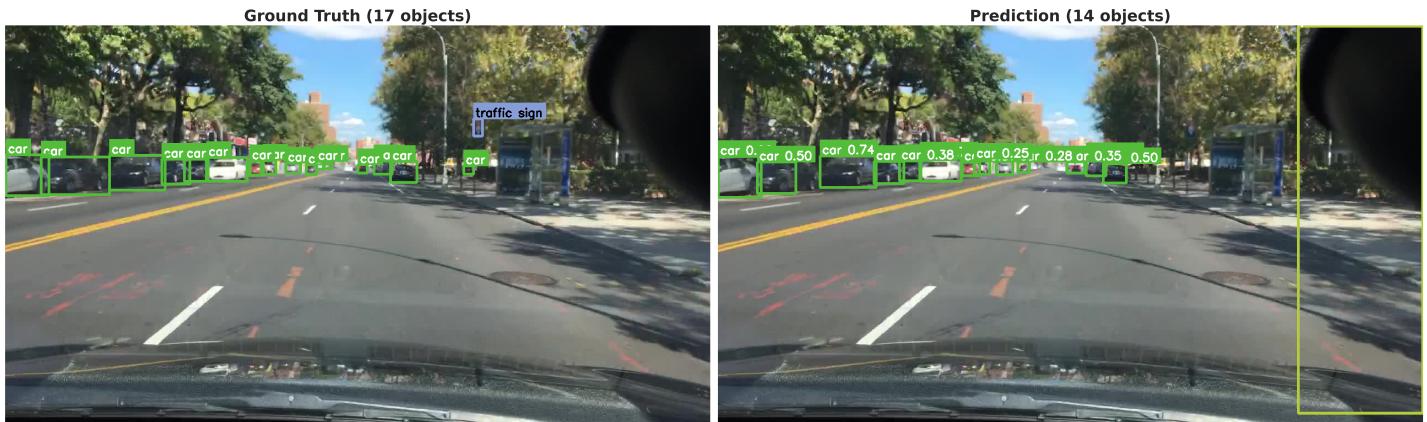




# Sample Predictions: Ground Truth vs Model

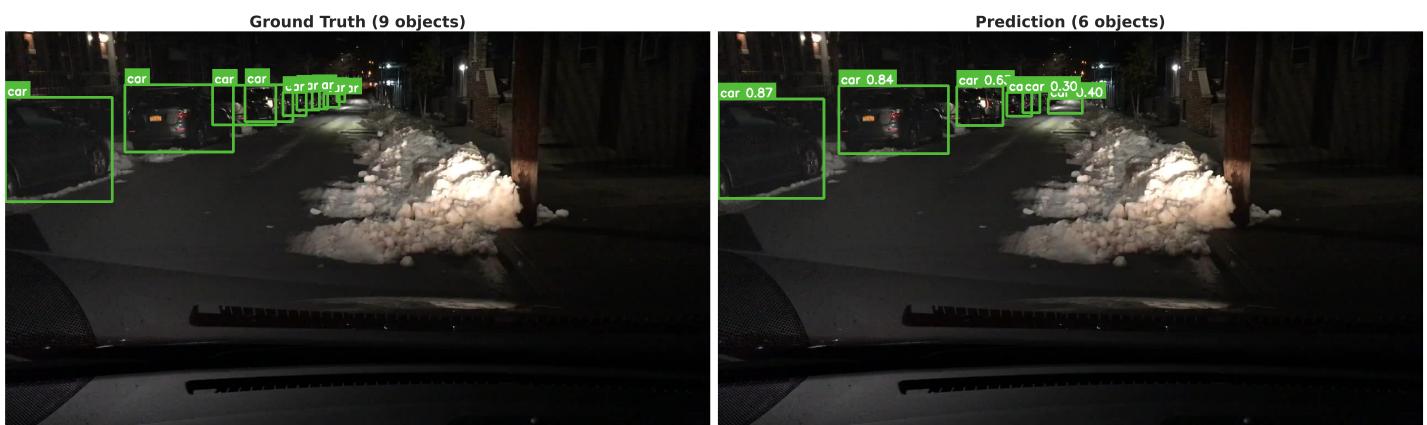
## Sample #1: d1349ed8-1bb98cfe.jpg

Ground Truth Objects: 17 | Predicted Objects: 14  
Weather: clear | Scene: highway | Time of Day: daytime



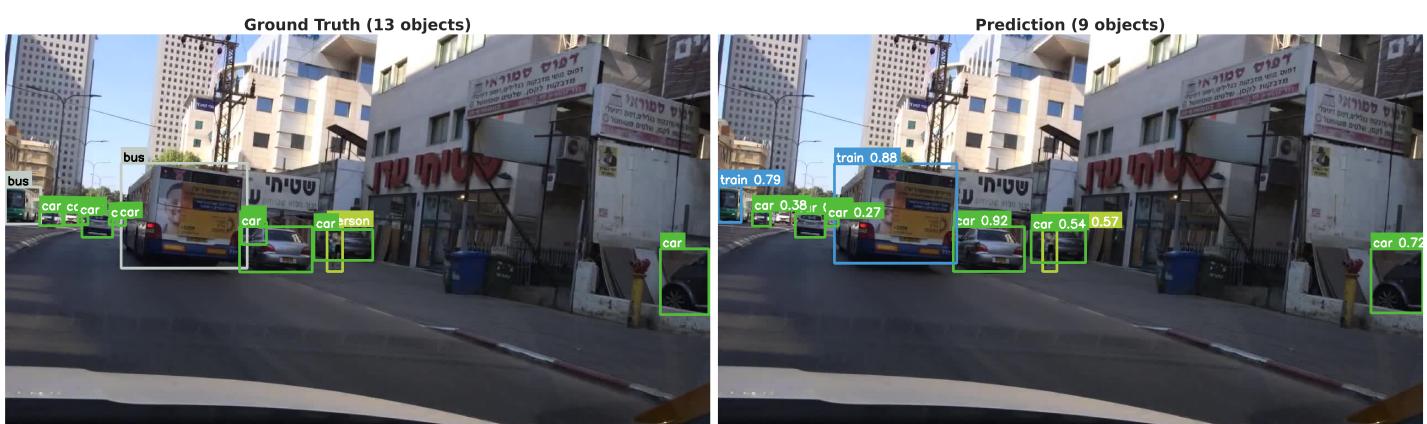
## Sample #2: ec209815-4dd80ba6.jpg

Ground Truth Objects: 9 | Predicted Objects: 6  
Weather: snowy | Scene: city street | Time of Day: night



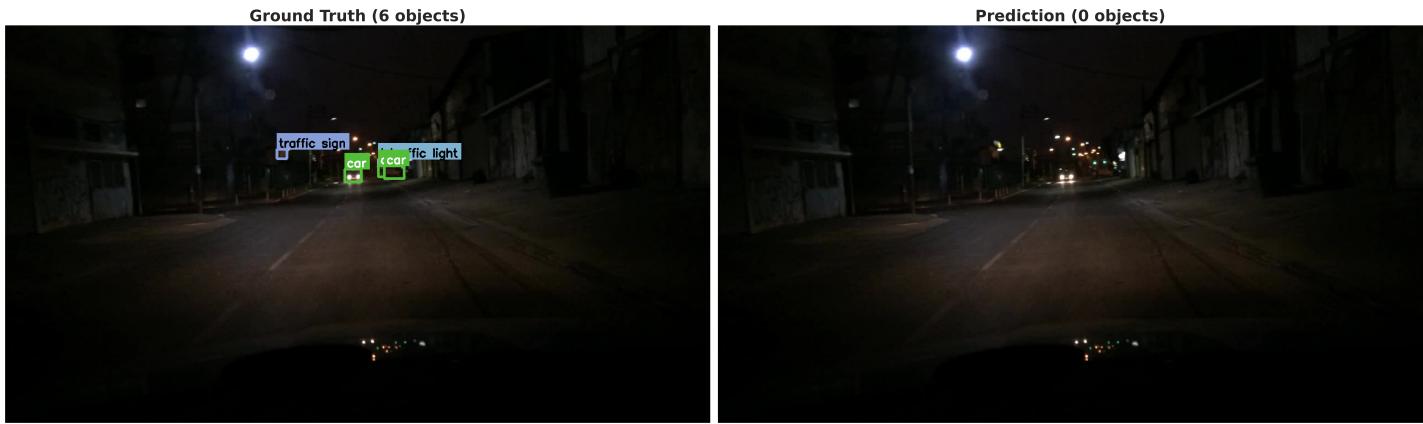
## Sample #3: d24b6aea-5a127b57.jpg

Ground Truth Objects: 13 | Predicted Objects: 9  
Weather: clear | Scene: city street | Time of Day: daytime



## Sample #4: e4cd8bdc-87096d65.jpg

Ground Truth Objects: 6 | Predicted Objects: 0  
Weather: clear | Scene: residential | Time of Day: night



**Sample #5: f6b20b28-e8ca5f90.jpg**

Ground Truth Objects: 18 | Predicted Objects: 11

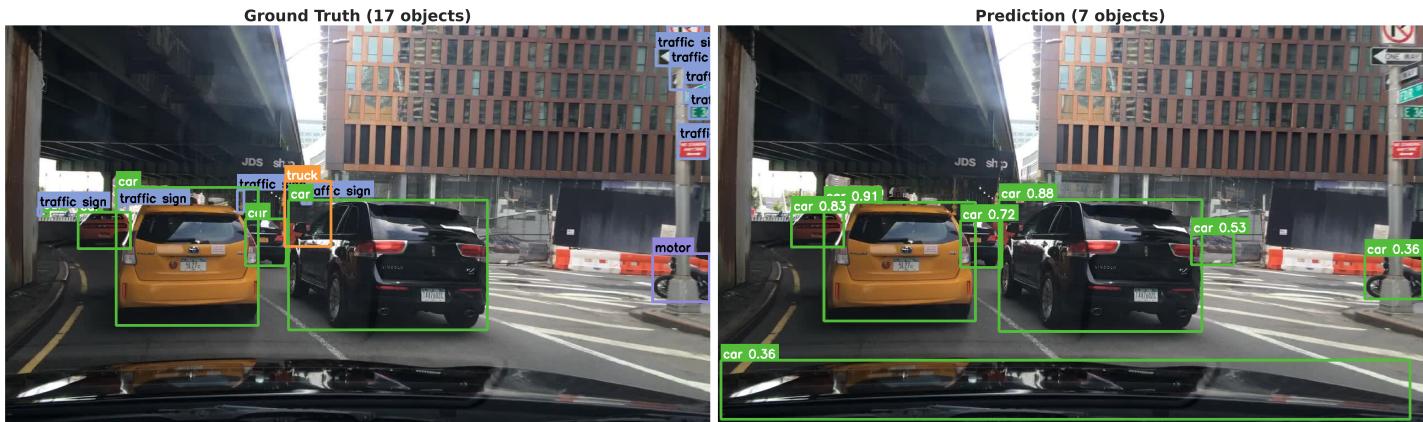
Weather: partly cloudy | Scene: city street | Time of Day: daytime



**Sample #6: f1adcfee-09451ed9.jpg**

Ground Truth Objects: 17 | Predicted Objects: 7

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



Generated by YOLO Testing Pipeline