

# YOLO Model Testing Report

<b>Model:</b>	yolov9m
<b>Model Size:</b>	39.1 MB
<b>Parameters:</b>	20.2 M
<b>FLOPs (640x640):</b>	77.87 GFLOPs
<b>Run Name:</b>	yolov9m_testing_20251127_223538
<b>Timestamp:</b>	2025-11-27 22:49:37
<b>Dataset:</b>	bdd100k_yolo_limited - test split
<b>Images Processed:</b>	20000
<b>Total Execution Time</b>	119.08s
<b>IoU Threshold:</b>	0.5

## Inference Performance

Metric	Value
Average Inference Time	2.68 ms per image
FPS (Frames Per Second)	372.96

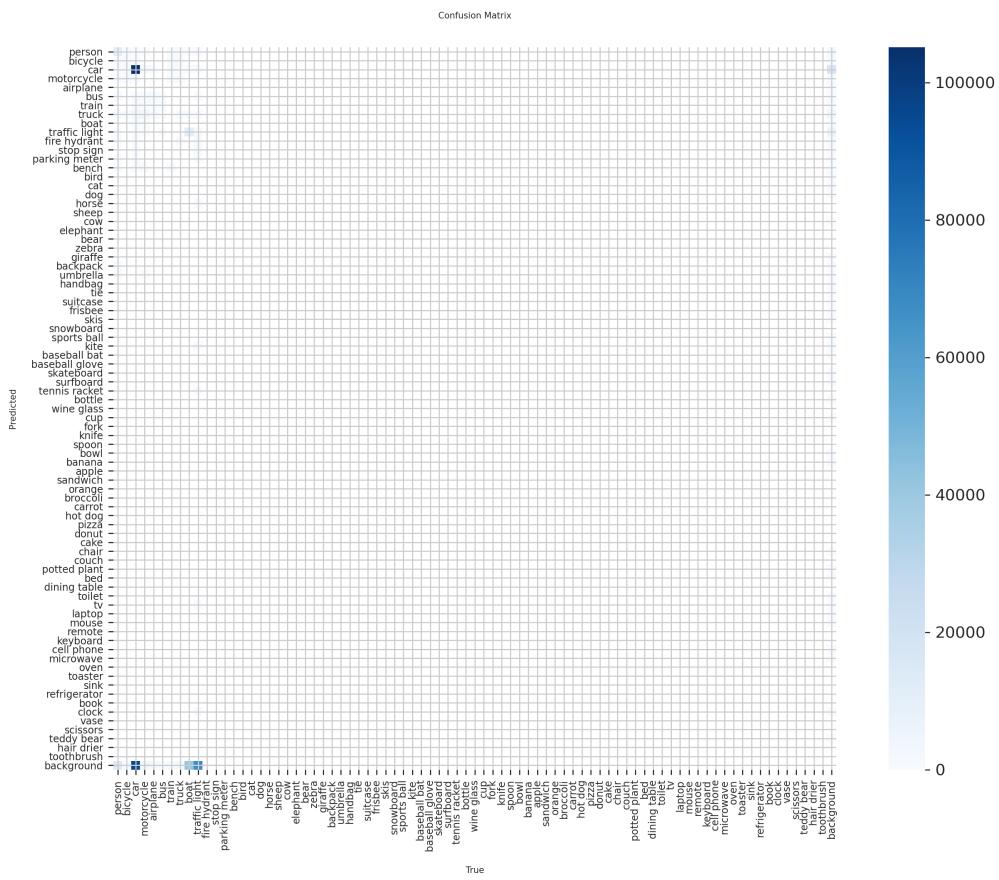
Note: Inference time includes preprocessing, model inference, and postprocessing.

## Overall Accuracy Metrics

Metric	Value
Precision	0.3167
Recall	0.6886
F1-Score	0.4339
mAP@0.5	0.1227
mAP@0.5:0.95	0.0671
True Positives	116467
False Positives	251260
False Negatives	52671

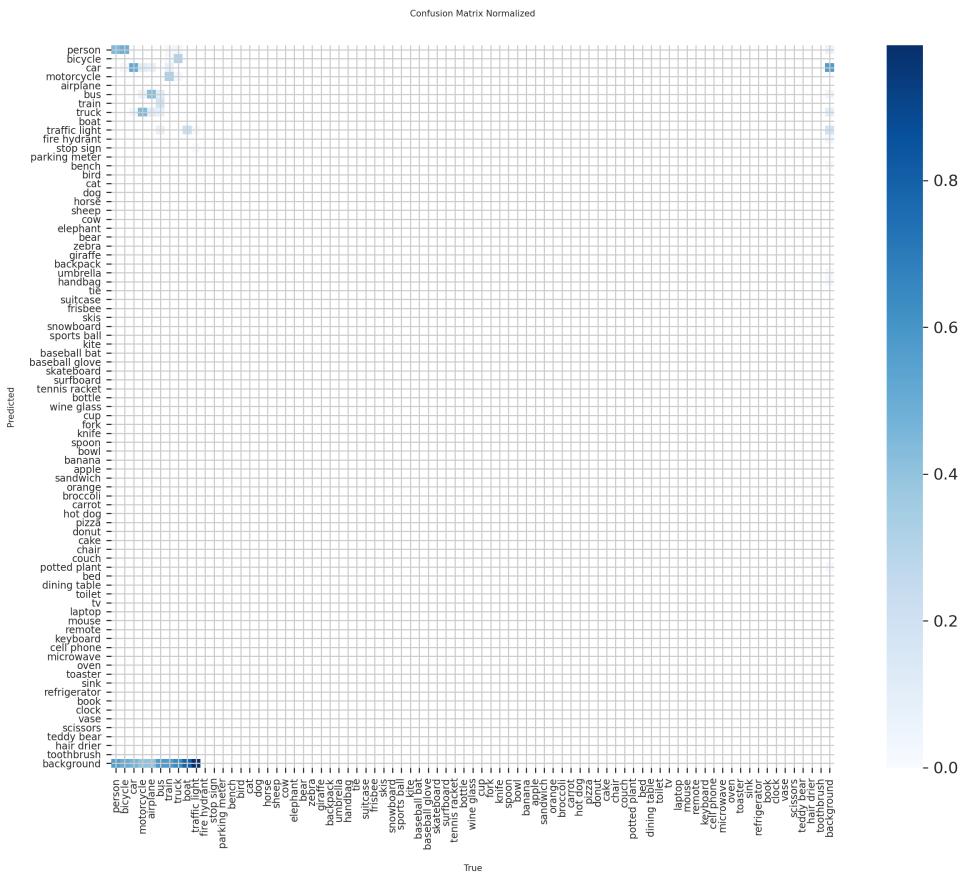
# Confusion Matrices

## Confusion Matrix (Counts)

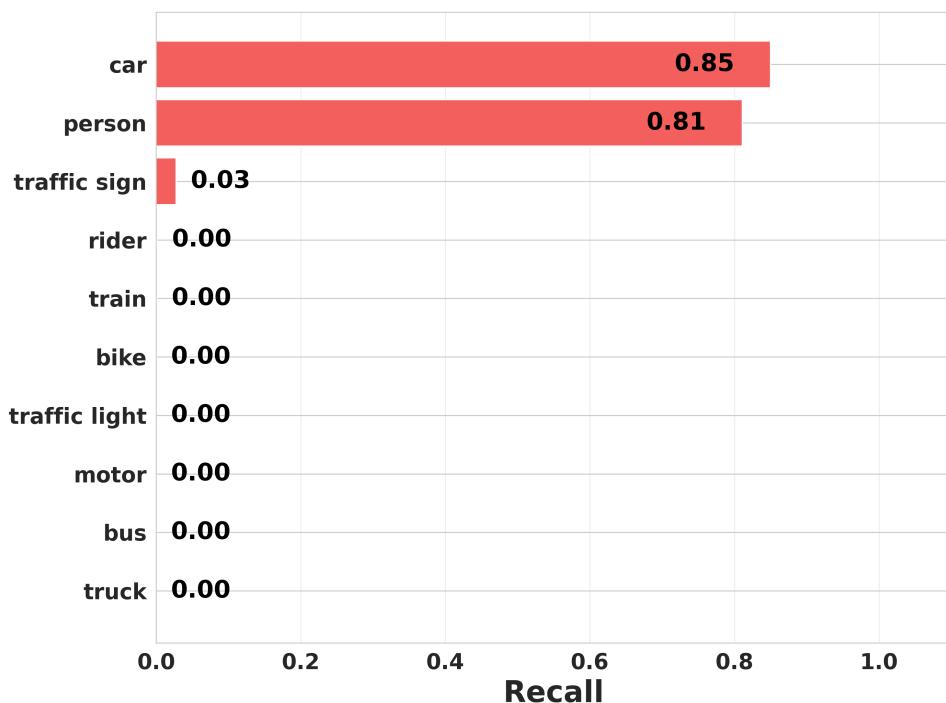


Note: Shows raw prediction counts for each class combination.

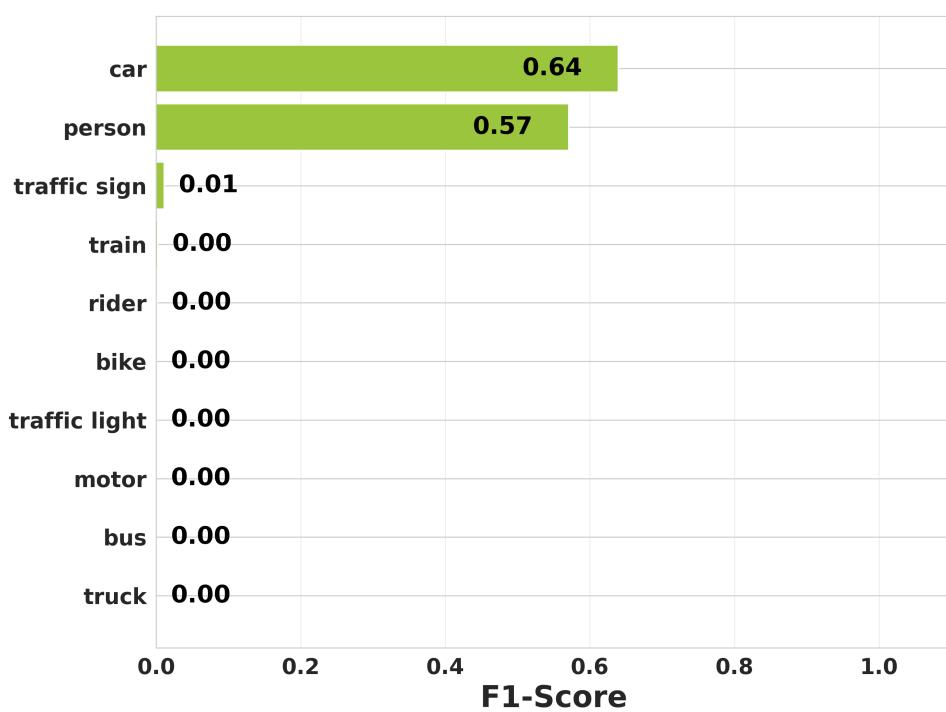
## Confusion Matrix (Normalized)



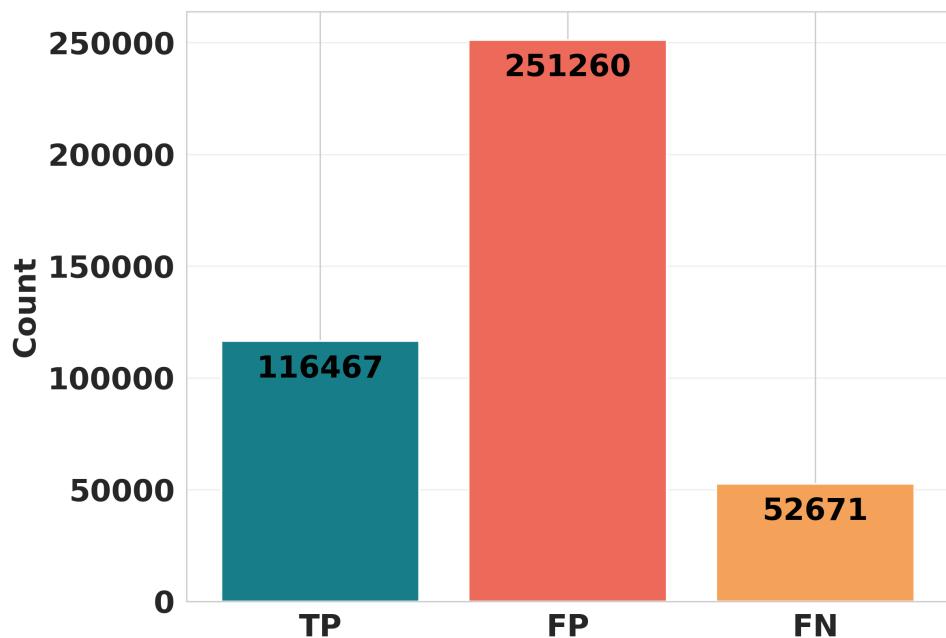
### Recall by Class



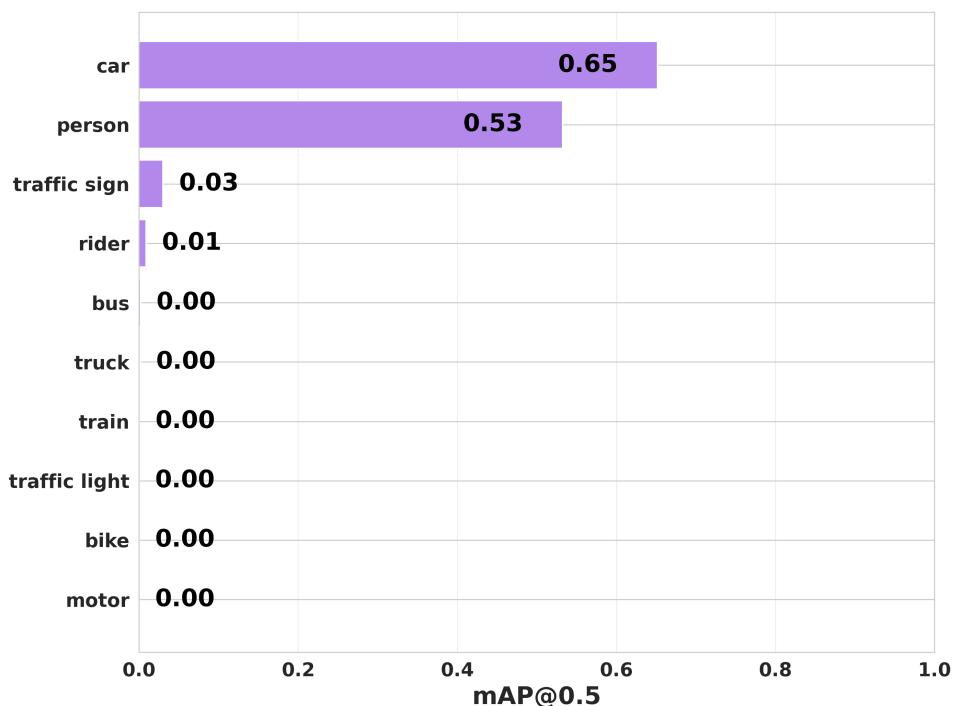
### F1-Score by Class



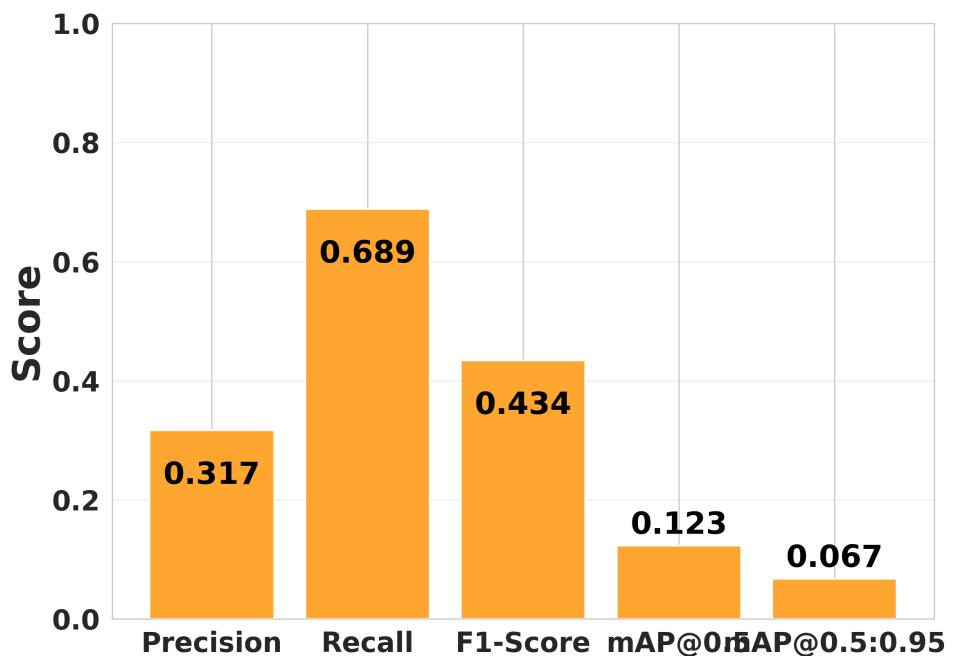
## Overall Detection Outcomes



## mAP@0.5 by Class



## Overall Metrics



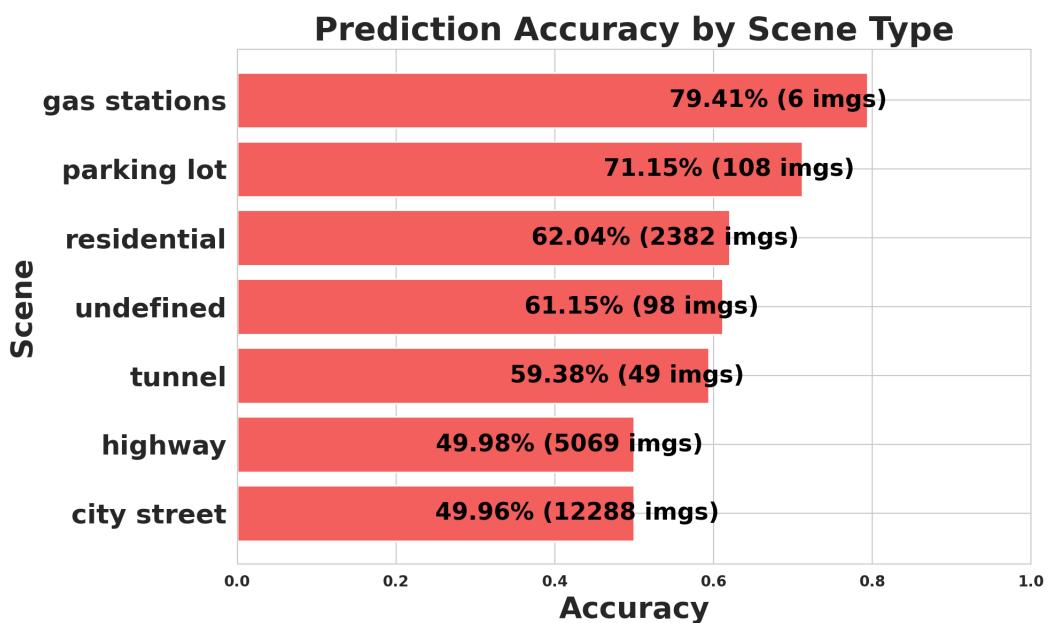
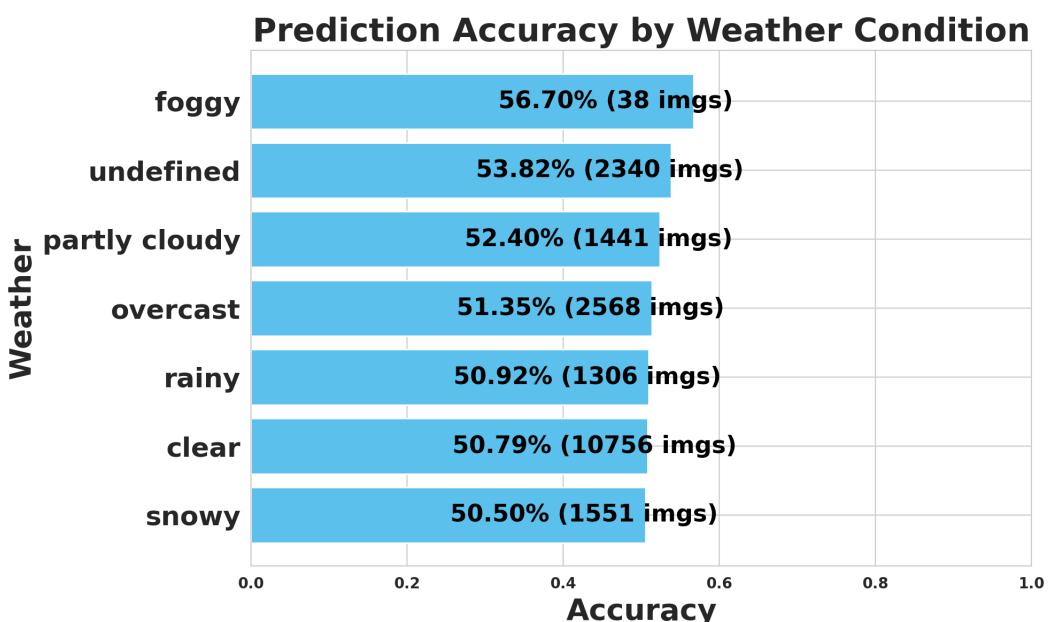
## Per-Class Performance

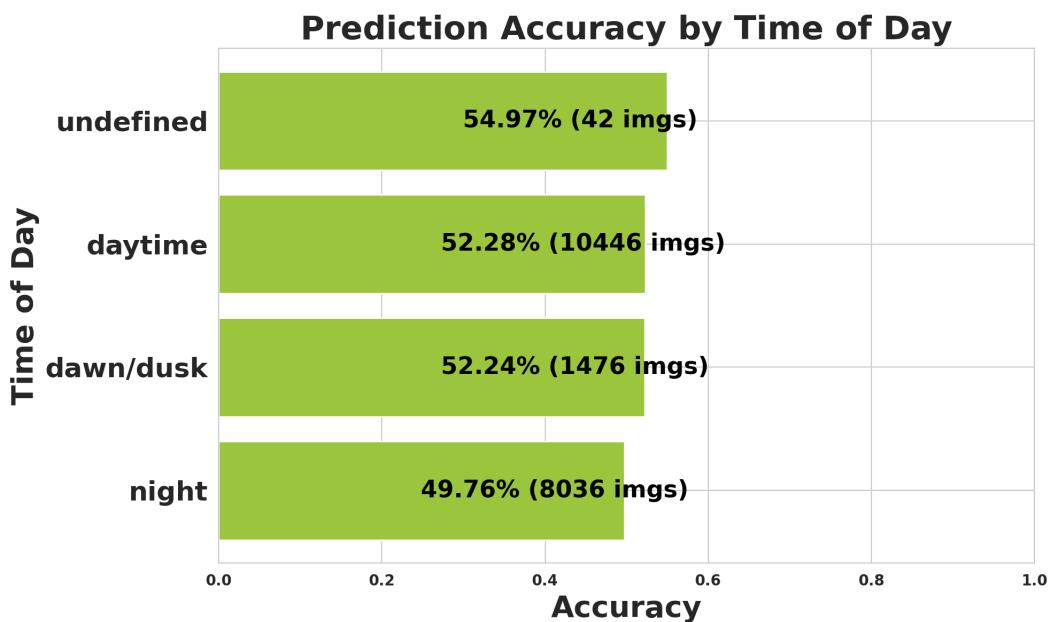
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	10860	13790	2535	0.4406	0.8108	0.5709	0.5326
rider	1	1293	723	0.0008	0.0014	0.0010	0.0087
car	105126	100023	18550	0.5124	0.8500	0.6394	0.6517
truck	0	8704	434	0.0000	0.0000	0.0000	0.0013
bus	0	3217	18	0.0000	0.0000	0.0000	0.0017
train	3	25	3014	0.1071	0.0010	0.0020	0.0003
motor	0	841	147	0.0000	0.0000	0.0000	0.0000
bike	1	1997	10419	0.0005	0.0001	0.0002	0.0001
traffic light	0	52840	12	0.0000	0.0000	0.0000	0.0001
traffic sign	476	68530	16819	0.0069	0.0275	0.0110	0.0302

# Comprehensive Failure Analysis

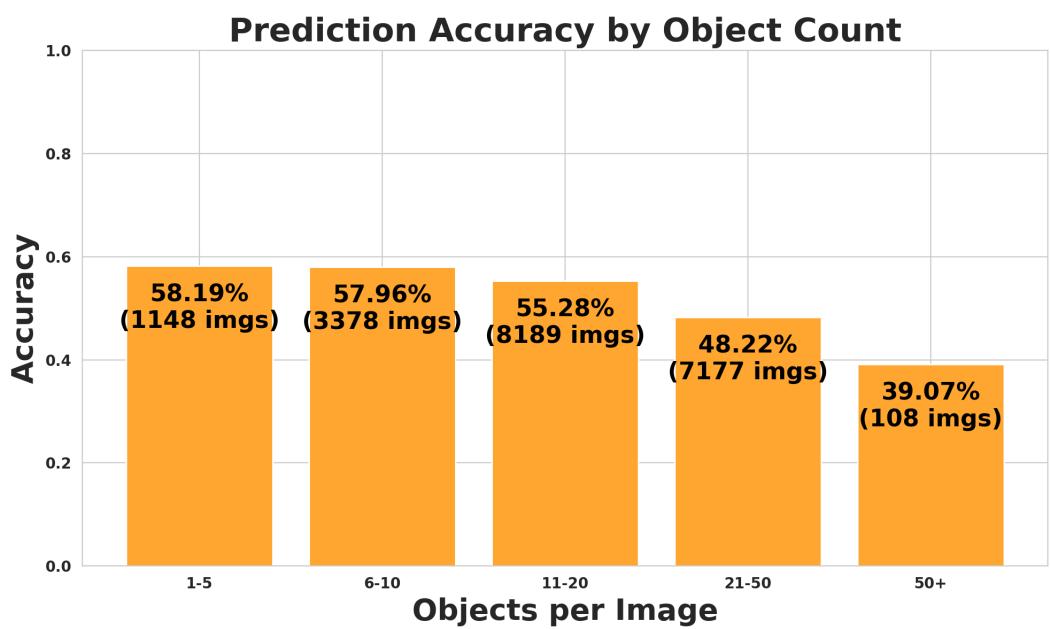
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	188925
Overall Accuracy	51.38%

## Accuracy by Environmental Attributes

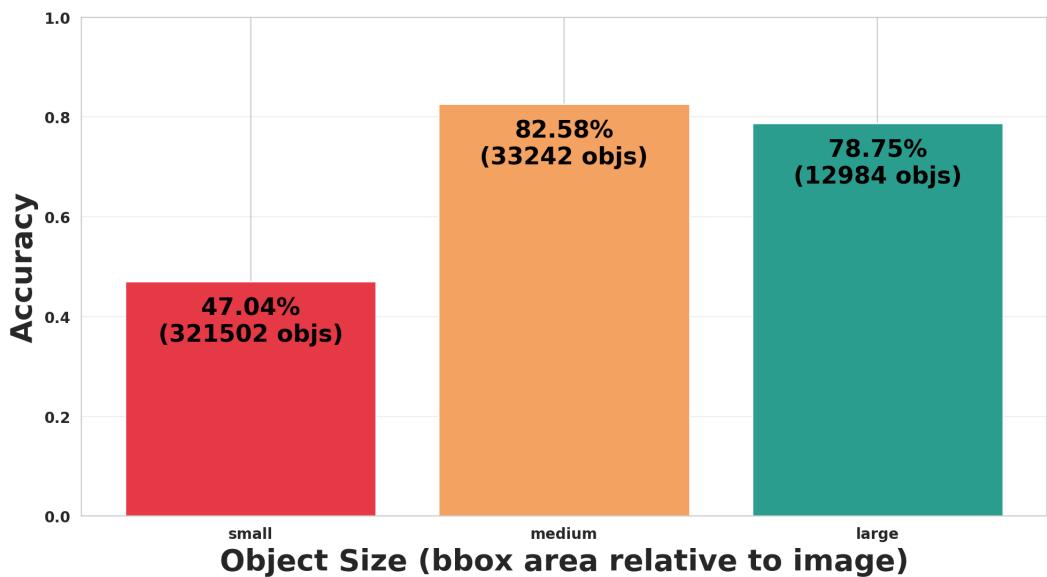




## Accuracy by Object Characteristics

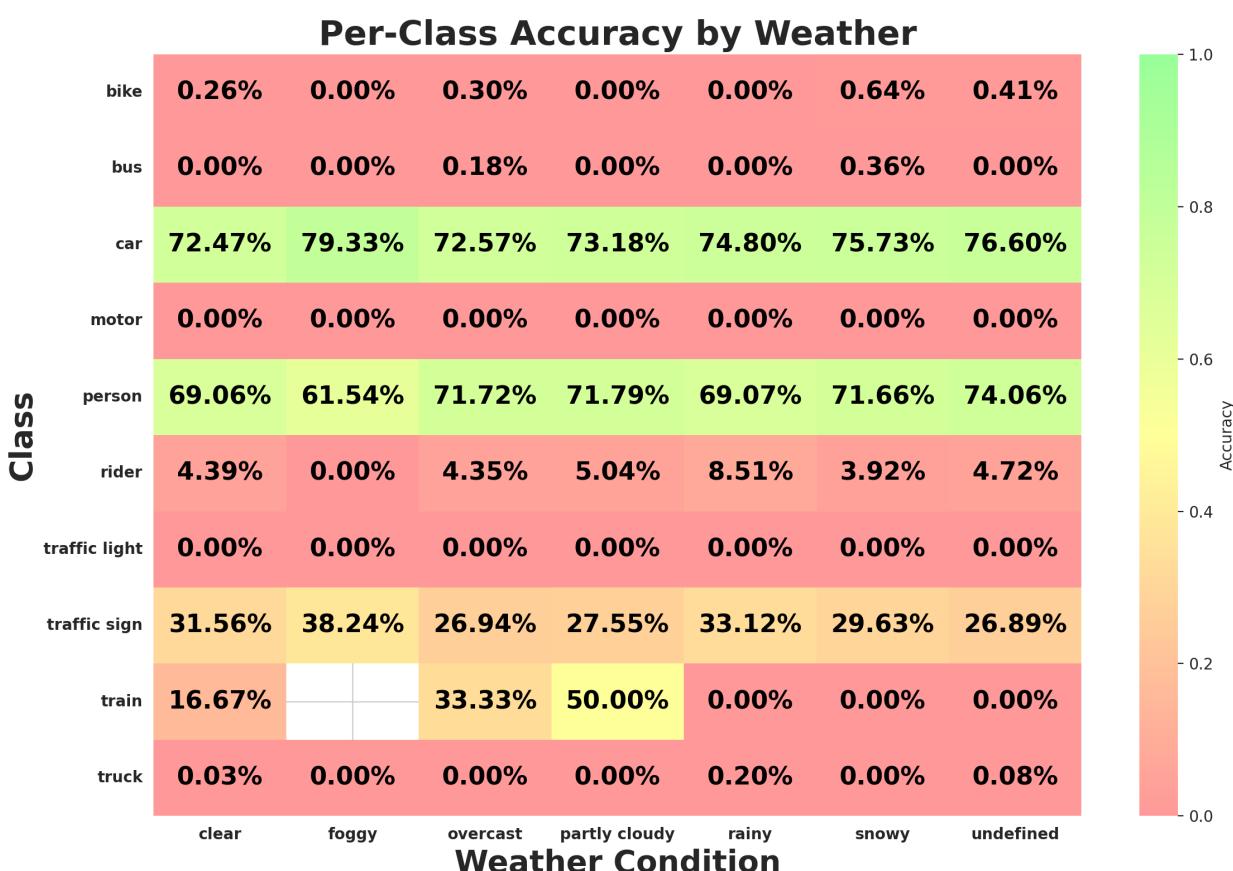
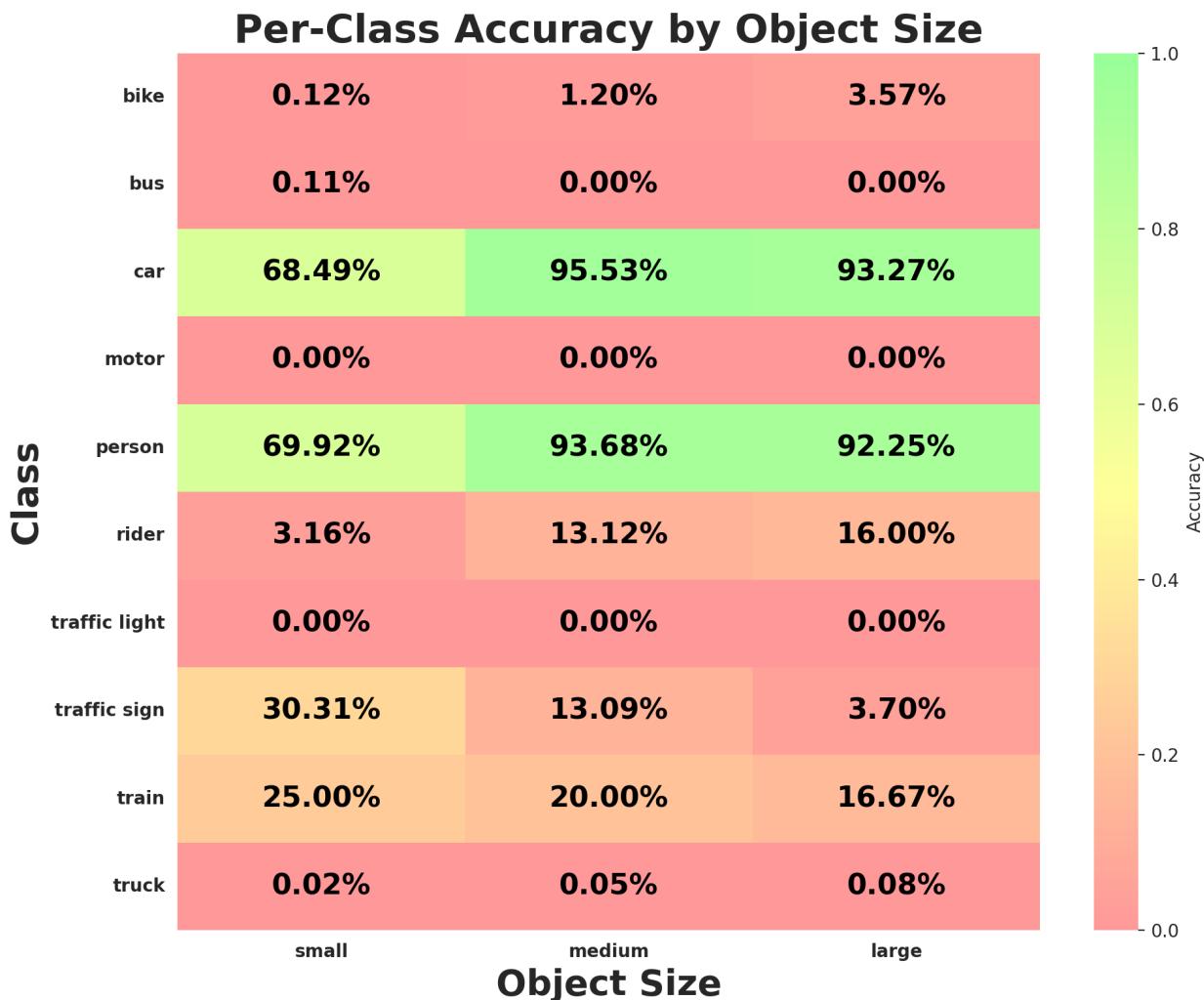


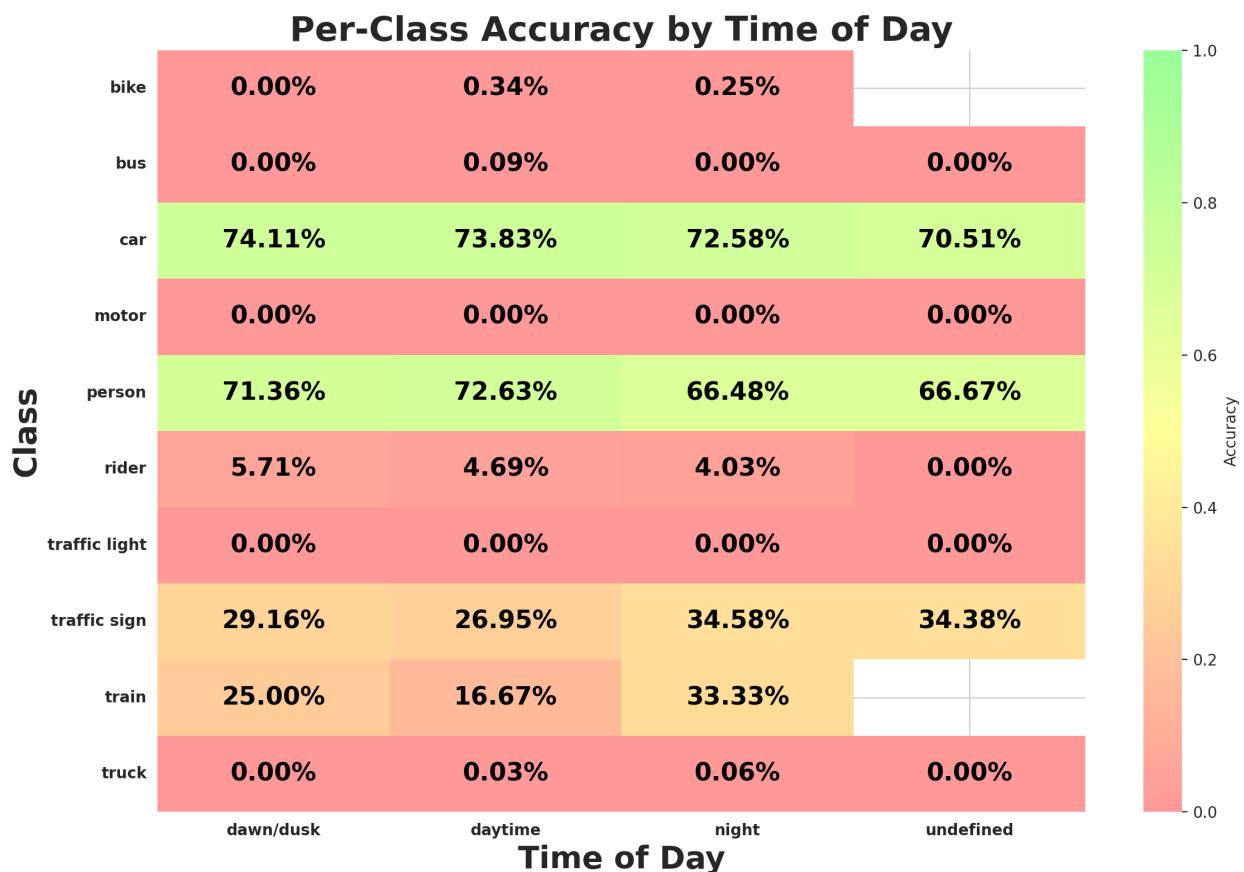
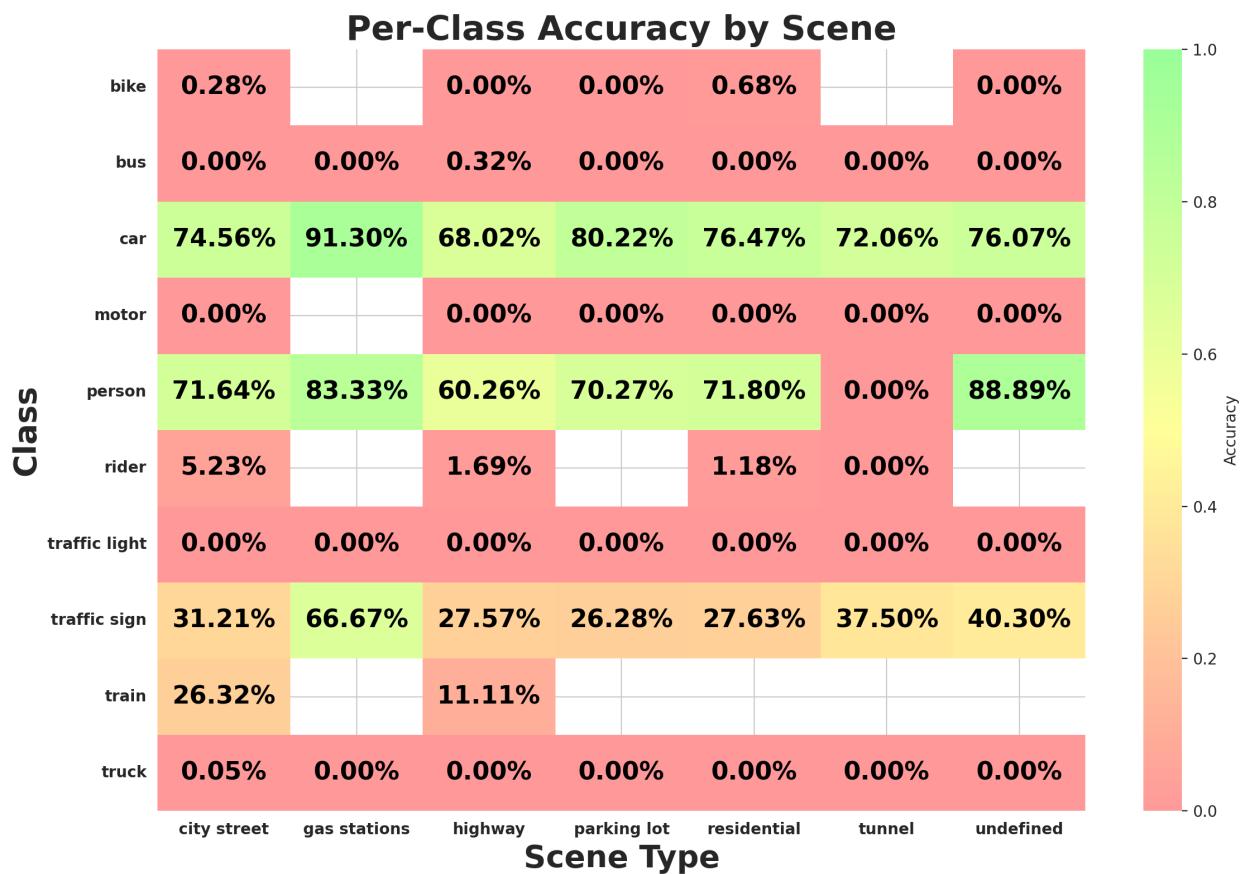
### Prediction Accuracy by Object Size



**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

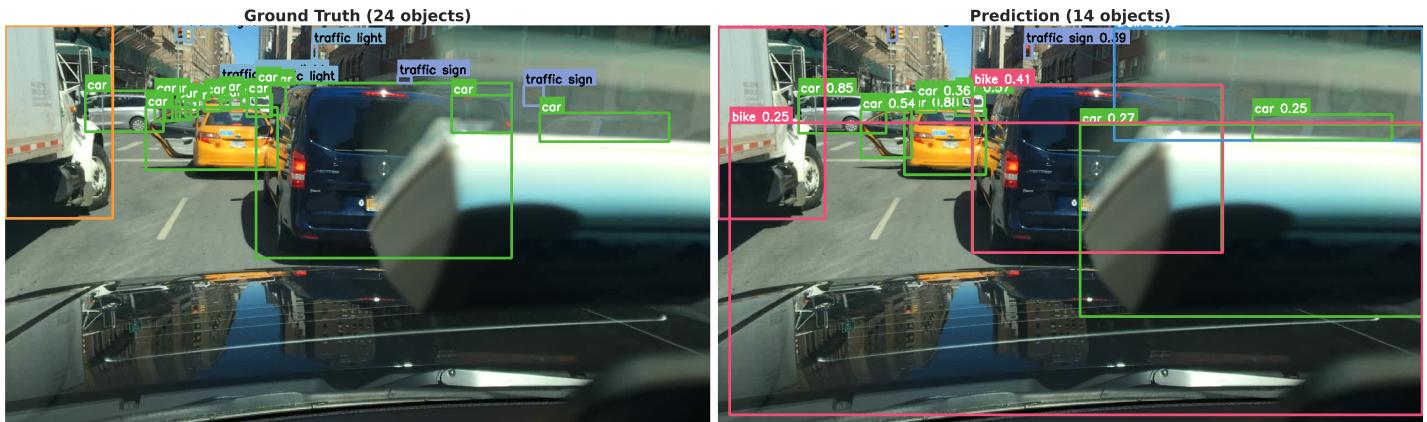




# Sample Predictions: Ground Truth vs Model

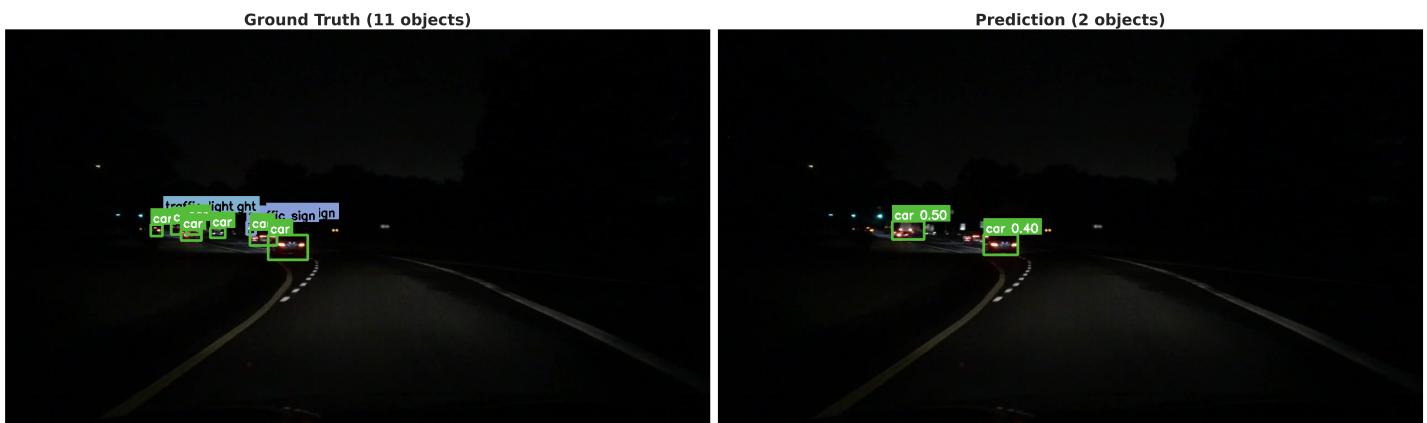
## Sample #1: f5ec9202-cf75689f.jpg

Ground Truth Objects: 24 | Predicted Objects: 14  
Weather: clear | Scene: city street | Time of Day: daytime



## Sample #2: f67dc8c0-b3b691cb.jpg

Ground Truth Objects: 11 | Predicted Objects: 2  
Weather: clear | Scene: highway | Time of Day: night



## Sample #3: da530bf7-69b7a4a0.jpg

Ground Truth Objects: 25 | Predicted Objects: 13  
Weather: clear | Scene: city street | Time of Day: daytime



## Sample #4: dc210de9-219eab9b.jpg

Ground Truth Objects: 22 | Predicted Objects: 8  
Weather: clear | Scene: highway | Time of Day: daytime



**Sample #5: d0d605e7-0c51ef58.jpg**

Ground Truth Objects: 10 | Predicted Objects: 8  
Weather: clear | Scene: city street | Time of Day: night



**Sample #6: f8d838c5-96d242bf.jpg**

Ground Truth Objects: 14 | Predicted Objects: 3  
Weather: overcast | Scene: highway | Time of Day: daytime



Generated by YOLO Testing Pipeline