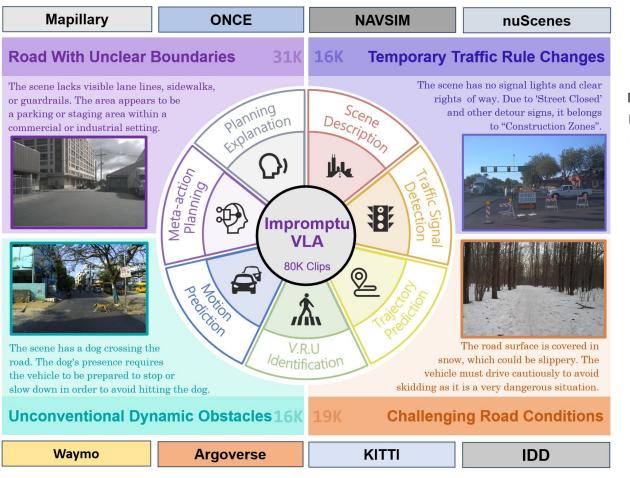
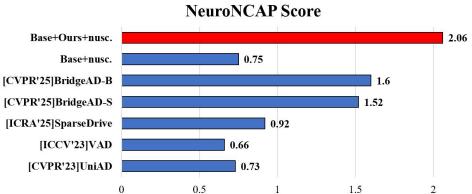
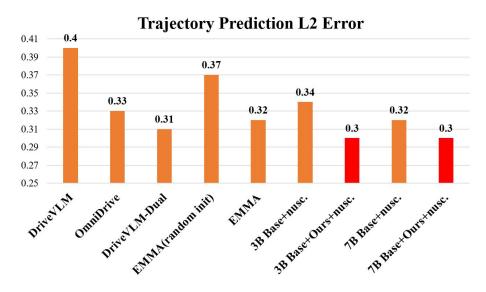
Impromptu VLA: Open Weights and Open Data for Driving Vision-Language-Action Models

Overview

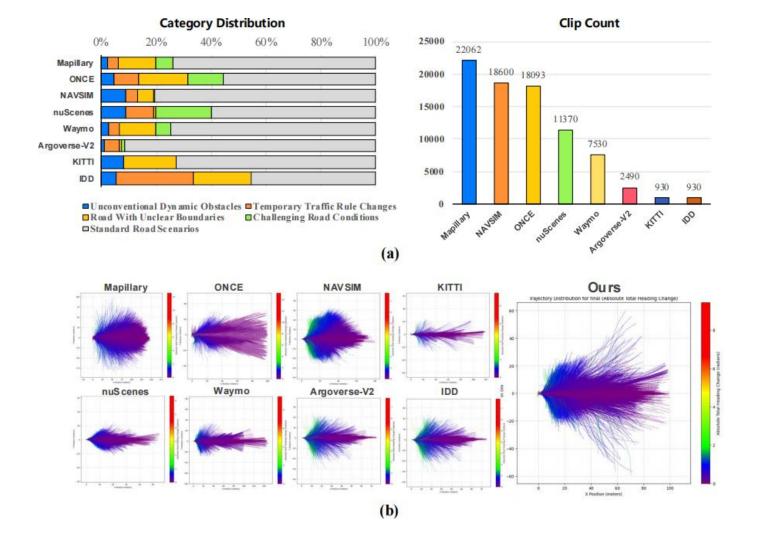
Impromptu Driving Vision-Language-Action Models





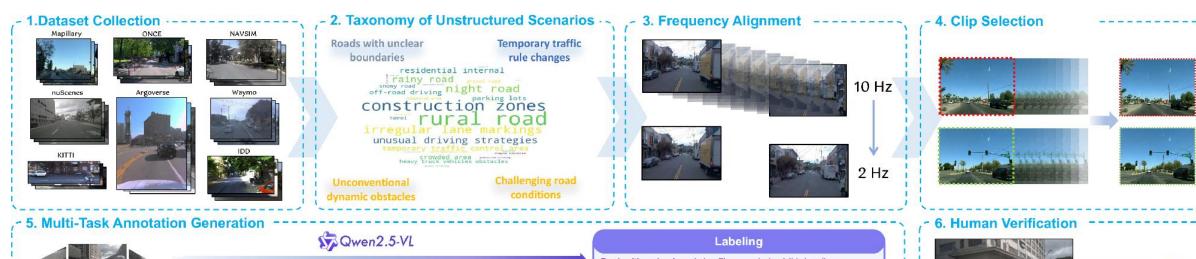


Impromptu VLA Dataset





Data Processing and Annotation Pipeline



Scene Description

- Time Period: afternoon
- Weather: cloudy - Scene Description:

in an open lot ...

The scene is set in an industrial or commercial area with multiple large trucks and trailers parked

Roadway Features

No visible lane lines are present, indicating no clear lane divisions for driving.

This may require the vehicle to navigate based on spatial awareness rather than following lane markings ...

Movable Objects

these obstacles carefully...

Multiple large trucks and trailers are parked in the area, both in front of and to the sides of the vehicle. These vehicles are stationary and pose no immediate driving hazard but limit the vehicle's maneuvering space. The vehicle must navigate around si

Roads with unclear boundaries: The scene lacks visible lane lines, sidewalks, or guardrails, fitting the Roads with unclear boundaries category. The secondary category is Residential internal roads as the area appears to be a parking or staging area within a commercial or industrial setting.

Temporary traffic rule changes: There are no construction zones, traffic cones, temporary signs or lights to indicate changes in traffic regulations. This category does not apply.

Unconventional dynamic obstacles: While there are stationary trucks, they are not dynamic obstacles. This category does not apply.

Challenging road conditions: The road surface is dry and solid, with no signs of slippery, muddy, gravel, snowy, icy, or unrecognizable conditions. This category does not apply.



Temporary traffic rule changes



Unconventional dynamic obstacle



Method	L2 Error (m) ↓			
	1s	2s	3s	Avg.
Ours and Key Competitors (Specie	alized Dr	iving M	odels)	
DriveVLM ³ [58]	0.18	0.34	0.68	0.40
OmniDrive ³ [61]	0.14	0.29	0.55	0.33
DriveVLM-Dual ³ [58]	0.15	0.29	0.48	0.31
EMMA (random init) [25] ³	0.15	0.33	0.63	0.37
EMMA $[25]^3$	0.14	0.29	0.54	0.32
$EMMA+^{3}[25]$	0.13	0.27	0.48	0.29
3B Base+nuScenes	0.14	0.30	0.58	0.34
3B Base+Impromptu+nuScenes	0.13	0.27	0.52	0.30
7B Base+nuScenes	0.13	0.28	0.55	0.32
7B Base+Impromptu+nuScenes	0.13	0.27	0.51	0.30

Source	Method	NeuroNCAP Score ↑			Collision rate (%) \downarrow				
		Avg.	Stat.	Frontal	Side	Avg.	Stat.	Frontal	Side
CVPR 2023	UniAD ²	0.73	0.84	0.10	1.26	88.6	87.8	98.4	79.6
ICCV 2023	VAD^2	0.66	0.47	0.04	1.45	92.5	96.2	99.6	81.6
ICRA 2025	SparseDrive ¹	0.92	-	-	_	93.9	-	-	-
CVPR 2025	BridgeAD-S ¹	1.52	=	-	-	76.2	-	-	-
CVPR 2025	BridgeAD-B ¹	1.60	-	-	-	72.6	-	-	-
	Base+nuScenes	0.75	0.99	0.55	0.70	90.0	88.6	93.2	88.0
-	Base+Impromptu+nuScenes	2.06	2.55	1.86	1.78	65.1	54.8	72.8	67.6

Conclusion

- Impromptu VLA Dataset address the critical data scarcity for autonomous driving in unstructured environments.
- Defines a taxonomy for unstructured driving scenarios
- Pushing forward the boundaries of existing end-to-end autonomous driving benchmarks in open-loop and closed-loop evaluation.
- We've released the code & dataset to encourage research on Impromptu VLA models for autonomous driving in unstructured scenarios.