AGC-Drive: A Large-Scale Dataset for Real-World Aerial-Ground Collaboration in Driving Scenarios

Yunhao Hou[1], Bochao Zou[1,*], Min Zhang[2], Ran Chen[1], Shangdong Yang[2], Junbao Zhuo[1], Siheng Chen[3], Jiansheng Chen[1], Huimin Ma[1,*]
[1]University of Science and Technology Beijing
[2]Xiamen NEVC Advanced Electric Powertrain Technology Innovation Center
[3]Shanghai Jiao Tong University

NEURAL INFORMATION PROCESSING SYSTEMS

Introduction

Collaborative perception cuts occlusions to boost driving accuracy. Aerial-Ground Collaborative Perception (AGCP) uses UAVs for top-down views that are easy to deploy, cost-effective, and flexible, enhancing blind spot coverage and long-range reasoning—vital for open roads and emergencies.

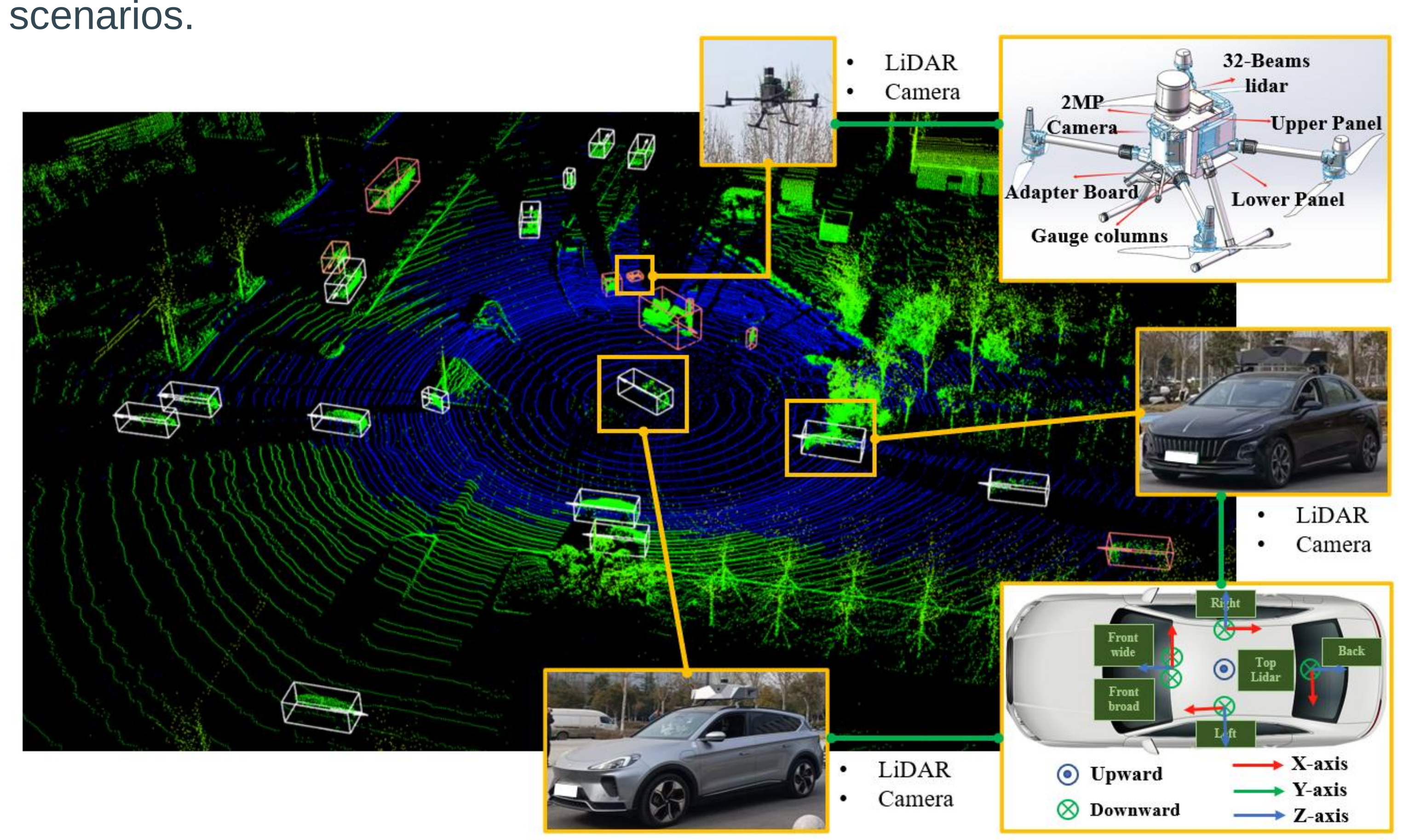
We built AGC-Drive, collecting data with two vehicles and one UAV. It offers:

- First **real-world** Air-Ground Coordination dataset for **driving scenario**.
- Large-scale, multi-modal, multi-view dataset across **14 scenario types**.
- UAV equipped with a **32-beam vehicle-grade LiDAR**.

Mode	Dataset	Year	Source	Agent	Sensor	scenario types	3D boxes	Classes	MvCams	Driving	UAV-L
V2V	OPV2V [1]	2022	Sim	Veh	C & L	6	230K	1	✓	✓	×
	V2V4Real [2]	2023	Real	Veh	C & L	-	240K	5	✓	✓	×
V2I	DAIR-V2X [6]	2022	Real	Veh & Inf	C & L	-	464K	10	×	✓	×
	V2X-Seq [7]	2023	Real	Veh & Inf	C & L	-	-	9	×	✓	×
	Rcooper [8]	2024	Real	Veh & Inf	C & L	-	-	10	×	✓	×
	TUMTraf-V2X [9]	2024	Real	Veh & Inf	C & L	-	29.3K	8	×	✓	×
	HoloVIC [10]	2024	Real	Veh & Inf	C & L	-	11.4M	3	×	✓	×
	V2X-R [11]	2025	Real	Veh & Inf	C & L & R	-	-	5	×	✓	×
V2V&I	V2X-Sim [3]	2022	Sim	Veh & Inf	C & L	-	26.6K	1	✓	✓	×
	V2XSet [4]	2022	Sim	Veh & Inf	C & L	5	230K	1	✓	✓	×
	V2X-Real [5]	2024	Real	Veh & Inf	C & L	-	1.2M	10	✓	✓	×
UAV	VisDrone [19]	2018	Real	UAV	C	-	10.2K	10	×	×	×
	UAVDT [20]	2018	Real	UAV	C	-	841.5K	3	✓	✓	×
U2U	CoPerception-UAV [13]	2023	Sim	UAV	C	-	1.6M	21	✓	✓	×
	UAV3D [14]	2023	Sim	UAV	C	-	3.3M	17	✓	✓	×
V2U	V2U-COO [17]	2024	Sim	Veh & UAV	C	-	-	4	×	✓	×
	CoPeD [16]	2024	Real	Veh & UAV	C & L	2	×	1	×	×	×
	Griffin [12]	2025	Sim	Veh & UAV	C & L	4	-	3	✓	✓	×
V2V&U	**AGC-Drive(Ours)**	**2025**	**Real**	**Veh & UAV**	**C & L & R**	**14**	**720K**	**13**	✓	✓	✓

System

AGC-Drive features a collaborative platform with 2 vehicles and 1 UAV. The vehicles are fitted with five cameras and one 128-beam LiDAR, while the UAV is equipped with one 32-beam LiDAR and a forward-facing camera, enabling comprehensive multi-view perception across diverse driving scenarios.

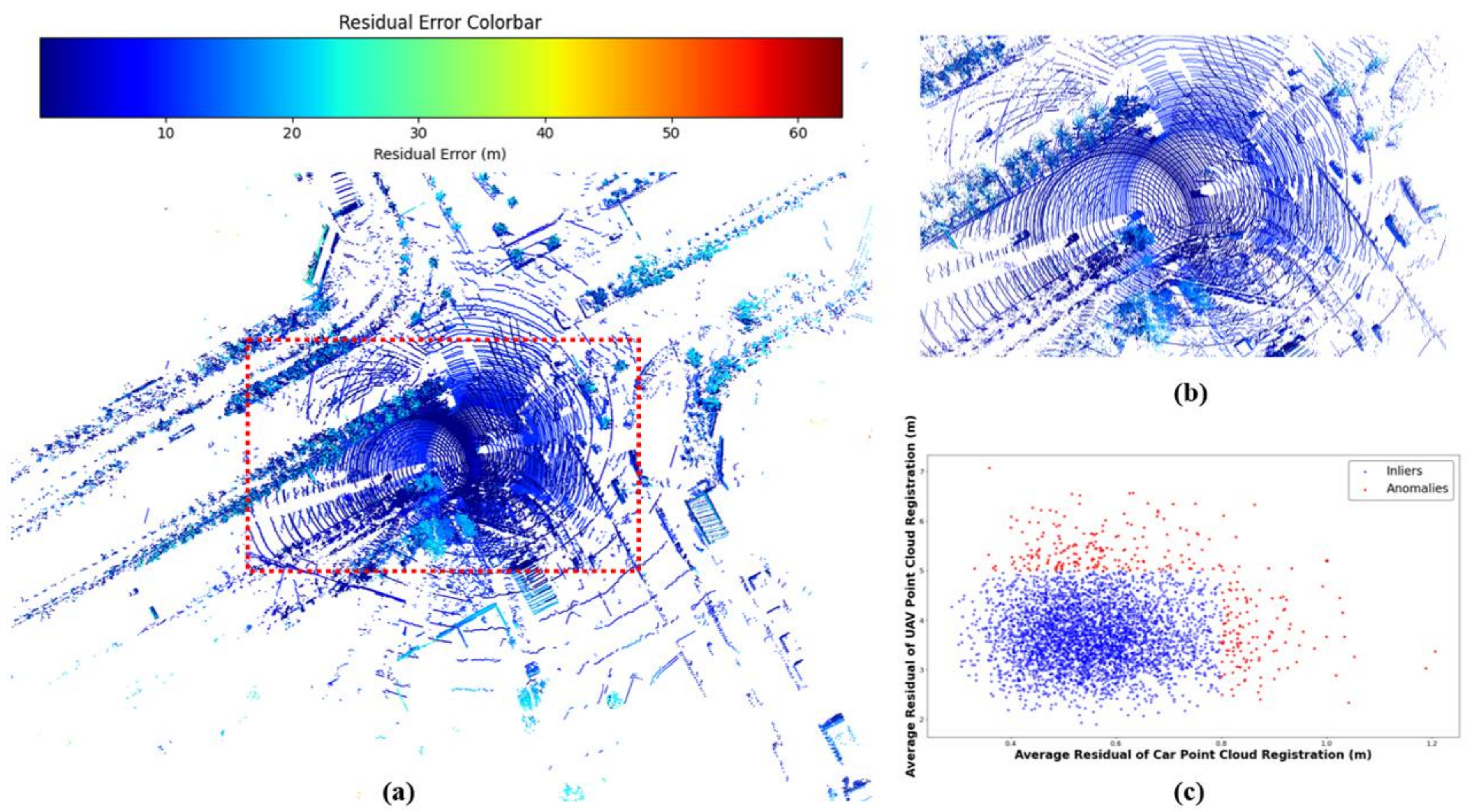


Collaborative data collection platform with 2 vehicles and 1 UAV.

Data Acquisition

Spatiotemporal Alignment

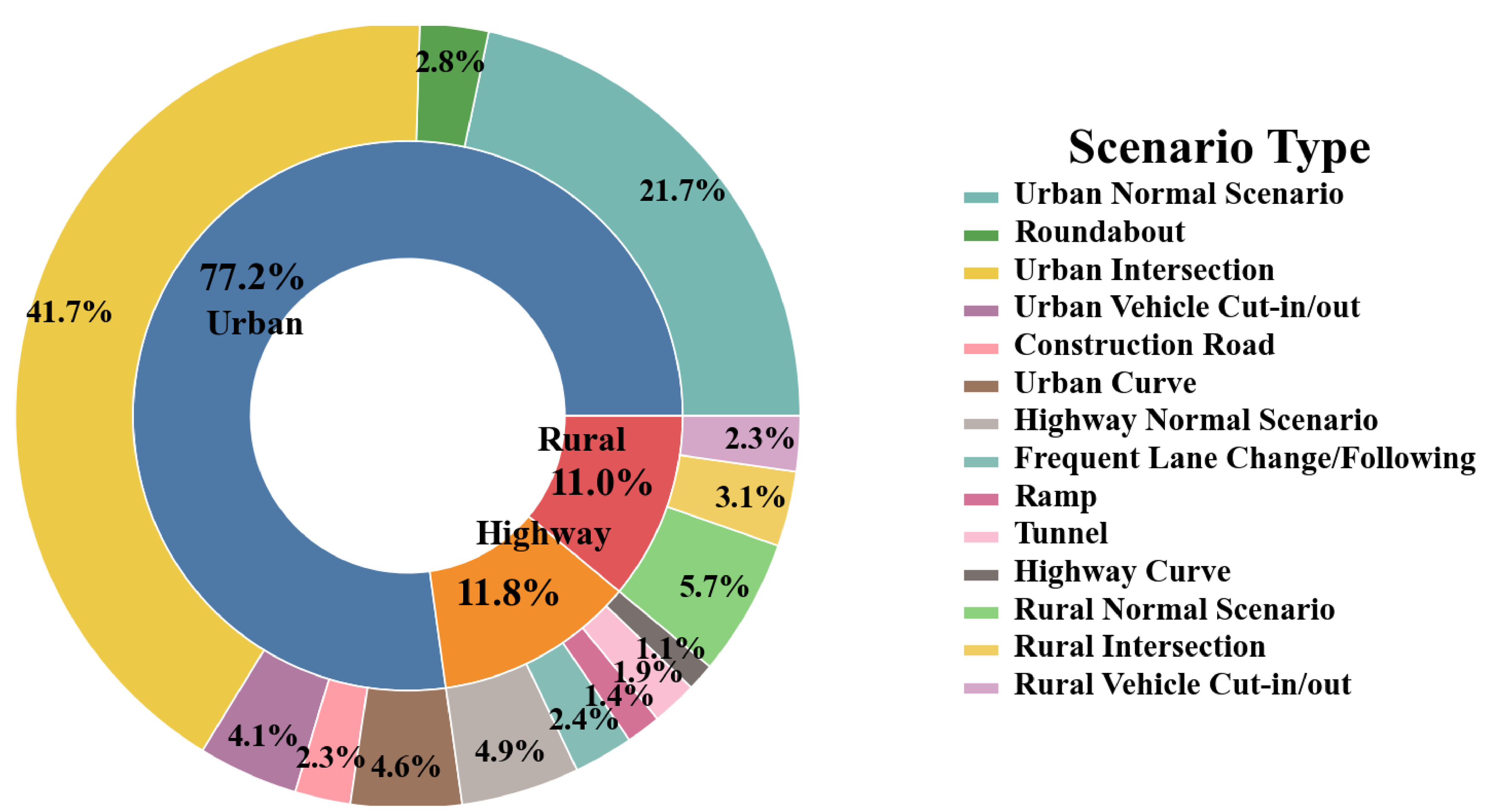
AGC-Drive ensures spatiotemporal alignment using unified GPS UTC timestamps for time synchronization and GPS/IMU data for initial ICP point cloud registration, followed by frame-by-frame manual refinement.



Sample error residuals from ICP point cloud registration.

Scenario Coverage

The dataset encompasses 14 diverse scenarios, including urban, rural, and highway environments, with 17% dynamic regions featuring vehicle cut-ins, cut-outs, and frequent lane changes.



Distribution of 14 scenario types in AGC-Drive.

Task And Benchmark

Benchmark for V2V 3D object detection

Table 3: 3D Detection Performance (%) on AGC-V2V.

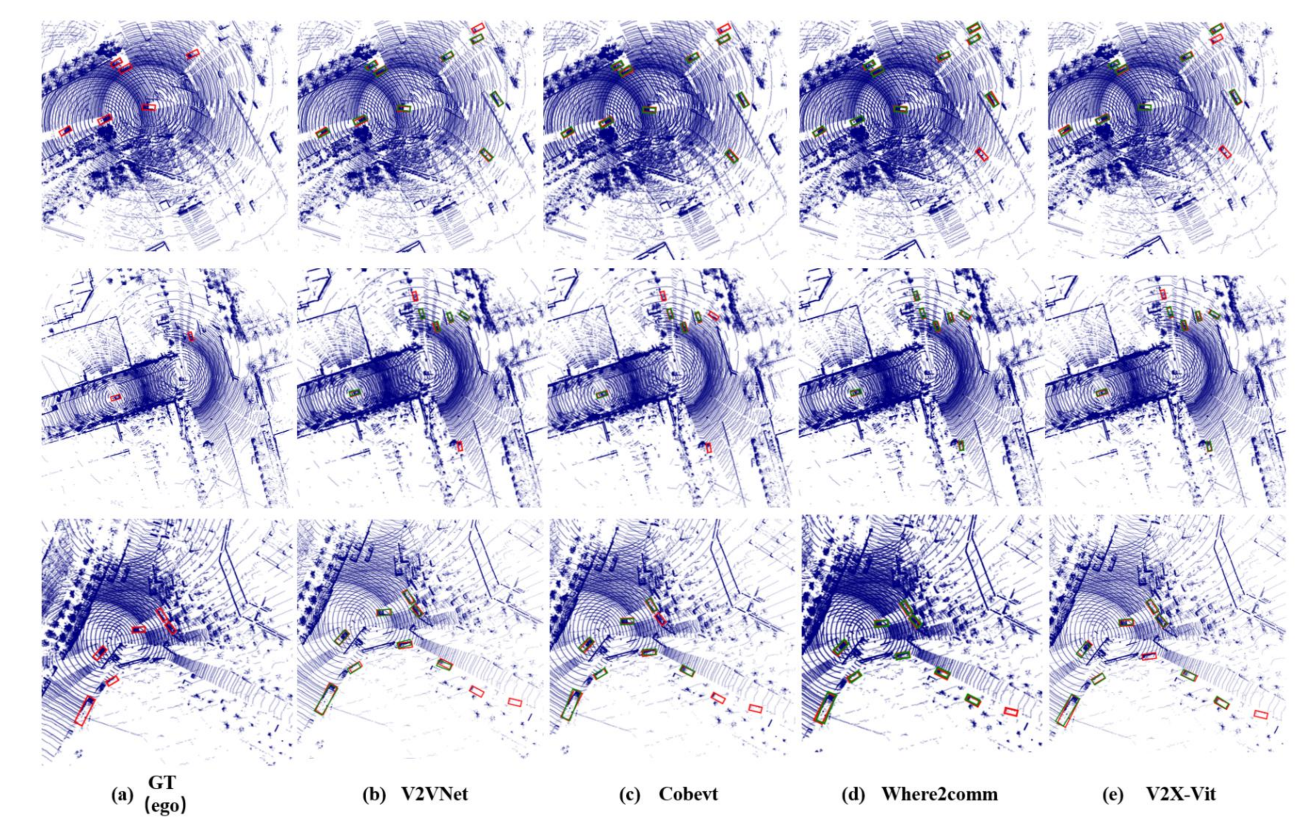
Co-Mode	Model	mAP@0.5	mAP@0.7
Late	PointPillars[27]	17.7	13.5
Early	PointPillars[27]	19.6	14.1
Intermediate	V2VNet [1]	18.4	5.7
	Cobevt [28]	**46.1**	**41.7**
	Where2comm [13]	39.3	31.5
	V2X-ViT [4]	44.1	36.6

Benchmark for VUC 3D object detection

Table 4: 3D Detection Performance (%) on AGC-VUC.

Co-Mode	Model	V2V		V2U		
		mAP@0.5	mAP@0.7	mAP@0.5	mAP@0.7	Δ_{UAV}
Intermediate	V2VNet [1]	30.5	14.6	**40.1**	**27.9**	**+11.5**
	Cobevt [28]	42.3	36.9	**42.9**	**37.5**	+0.6
	Where2comm [13]	42.6	30.7	**44.2**	**32.0**	+1.5
	V2X-ViT [4]	38.3	28.7	**42.6**	**33.9**	+4.8

visualization



References

Xu R, Xiang H, Xia X, et al.Opv2v: An open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication[C] //2022 International Conference on Robotics and Automation (ICRA). IEEE, 2022: 2583-2589.

Wang J, Cao X, Zhong J, et al. Griffin: Aerial-ground cooperative detection and tracking dataset and benchmark[J]. arXiv preprint arXiv:2503.06983, 2025.

Yue Hu, Shaoheng Fang, Zixing Lei, Yiqi Zhong, and Siheng Chen. Where2comm: Communication-efficient collaborativeperception via spatial confidence maps. Advances in neural information processing systems, 35:4874–4886, 2022.