NEURAL INFORMATION PROCESSING SYSTEMS

AGC-Drive: A Large-Scale Dataset for Real-World Aerial–Ground Collaboration

Yunhao Hou[1], Bochao Zou[1,*], Min Zhang[2], Ran Chen[1], Shangdong Yang[2], Yanmei Zhang[2], Junbao Zhuo[1], Siheng Chen[3], Jiansheng Chen[1], Huimin Ma[1,*]

北京科技大学 University of Science and Technology Beijing

厦门国创中心 NEVC-Xiamen

上海交通大学 SHANGHAI JIAO TONG UNIVERSITY

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Track on Datasets and Benchmarks

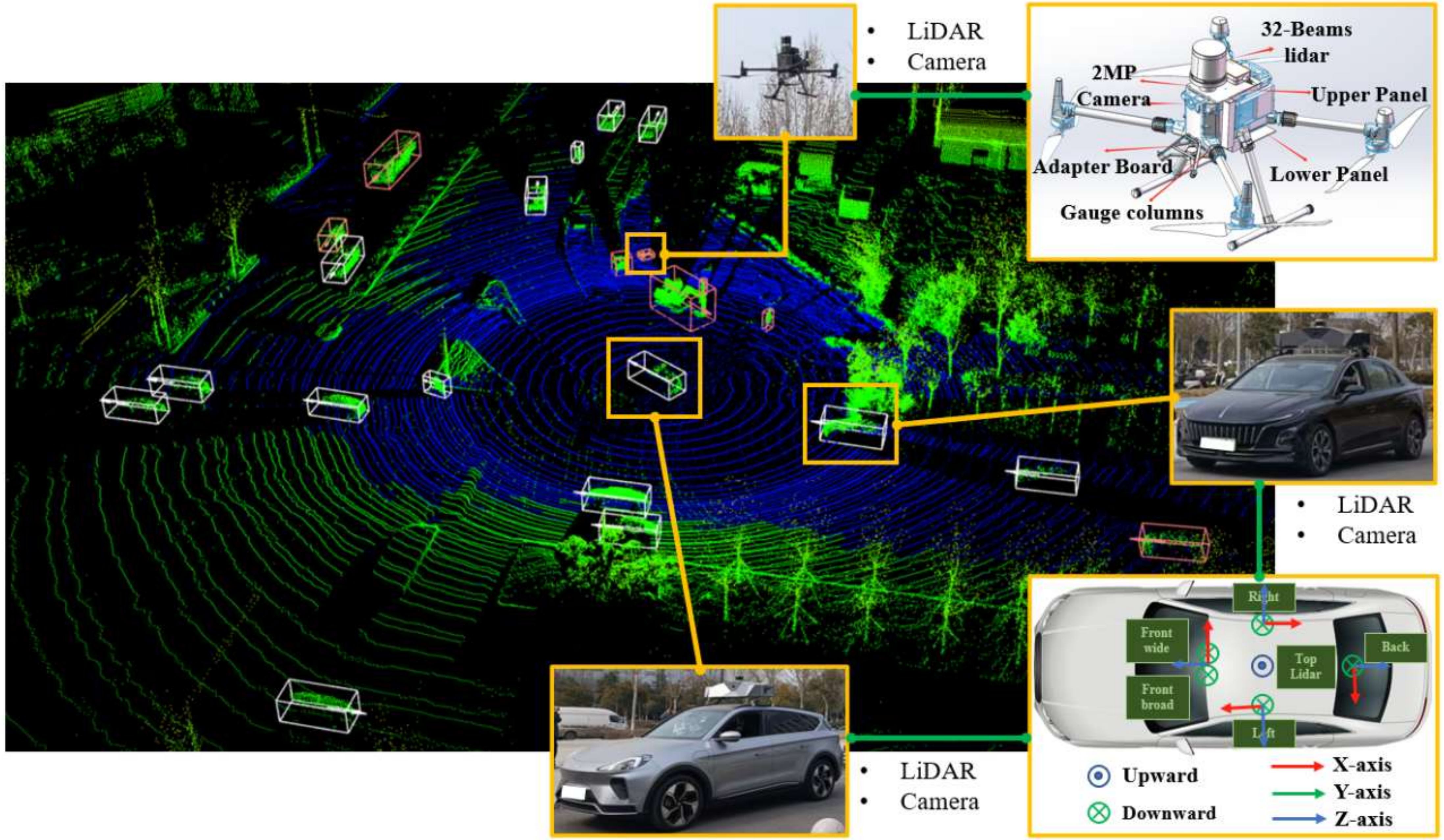
Background & Motivation (I)

◇ Autonomous driving perception is limited by **occlusions** and **restricted fields of view**.

◇ Collaborative perception across vehicles improves coverage but **remains ground-limited**.

◇ Most existing datasets lack **vertical diversity** and **aerial context**.

◇ Understanding cross-domain perception motivates our Aerial–Ground dataset design.



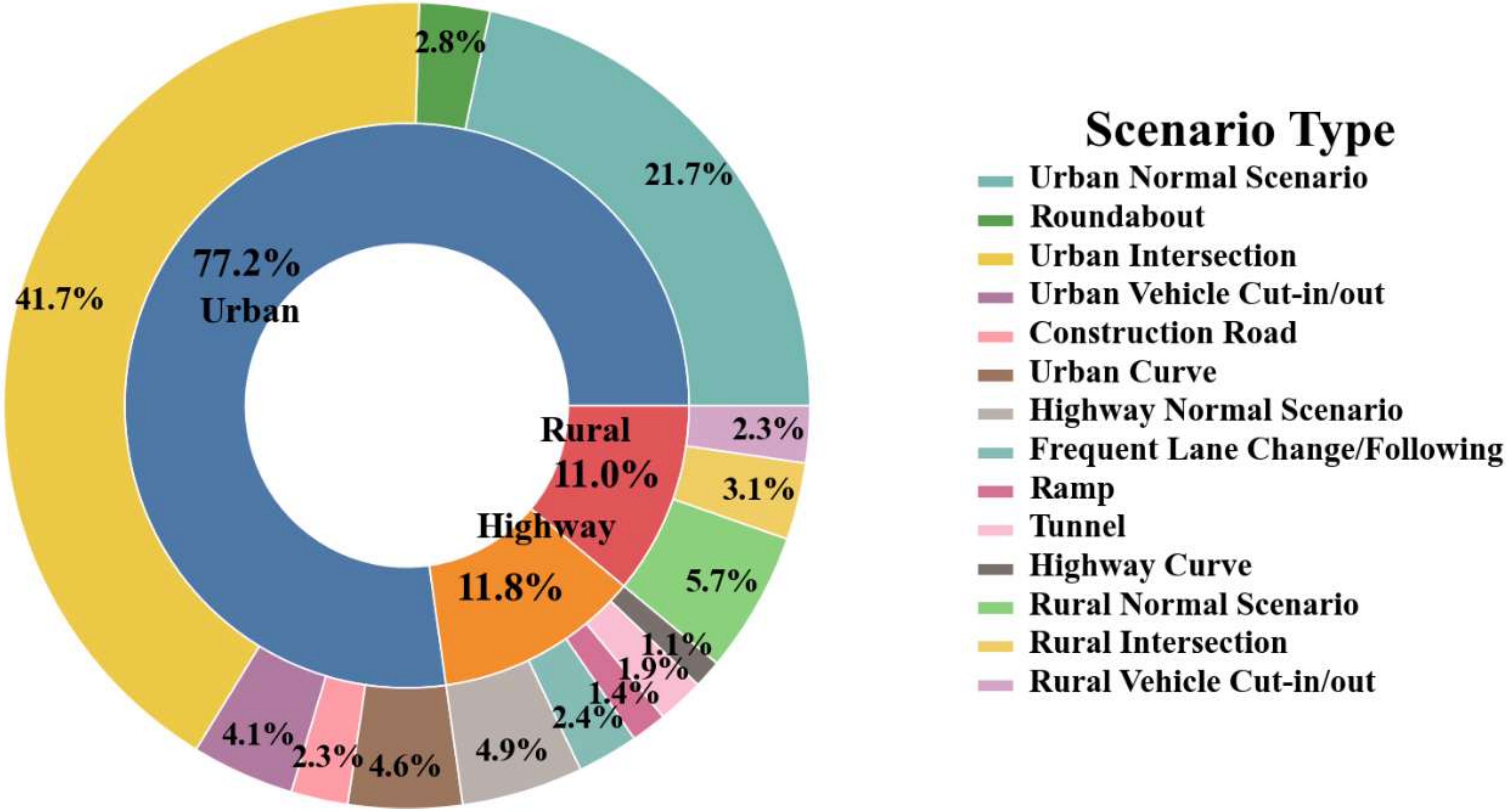
Background & Motivation (II)

- UAVs offer **dynamic, top-down** perspectives to complement vehicle sensors.
- Existing UAV datasets are **small, synthetic,** or **lack LiDAR** and calibration.
- **Real-world** UAV–vehicle cooperation remains underexplored.
- AGC-Drive bridges this gap with real-world, synchronized air-ground data.

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**	✓	✓	✓

Dataset Overview

- 80 hours of driving data covering **14** diverse real-world scenarios.
- **80K** LiDAR frames, **360K** images, and 720K annotated 3D bounding boxes.
- **13** object categories with occlusion-level annotations.
- 17% of scenes involve **dynamic events** such as cut-ins and merges.



System Setup

- Data collection platform: two vehicles + one UAV.
- Each vehicle: **128-beam LiDAR, 5 cameras, GPS/IMU**.
- UAV: **32-beam LiDAR, front camera, RTK GPS/IMU**.
- Sensors **synchronized** via GPS time; **calibrated** by GPS/IMU and ICP alignment.

Table 2: Key Sensor Specifications in AGC-Drive.

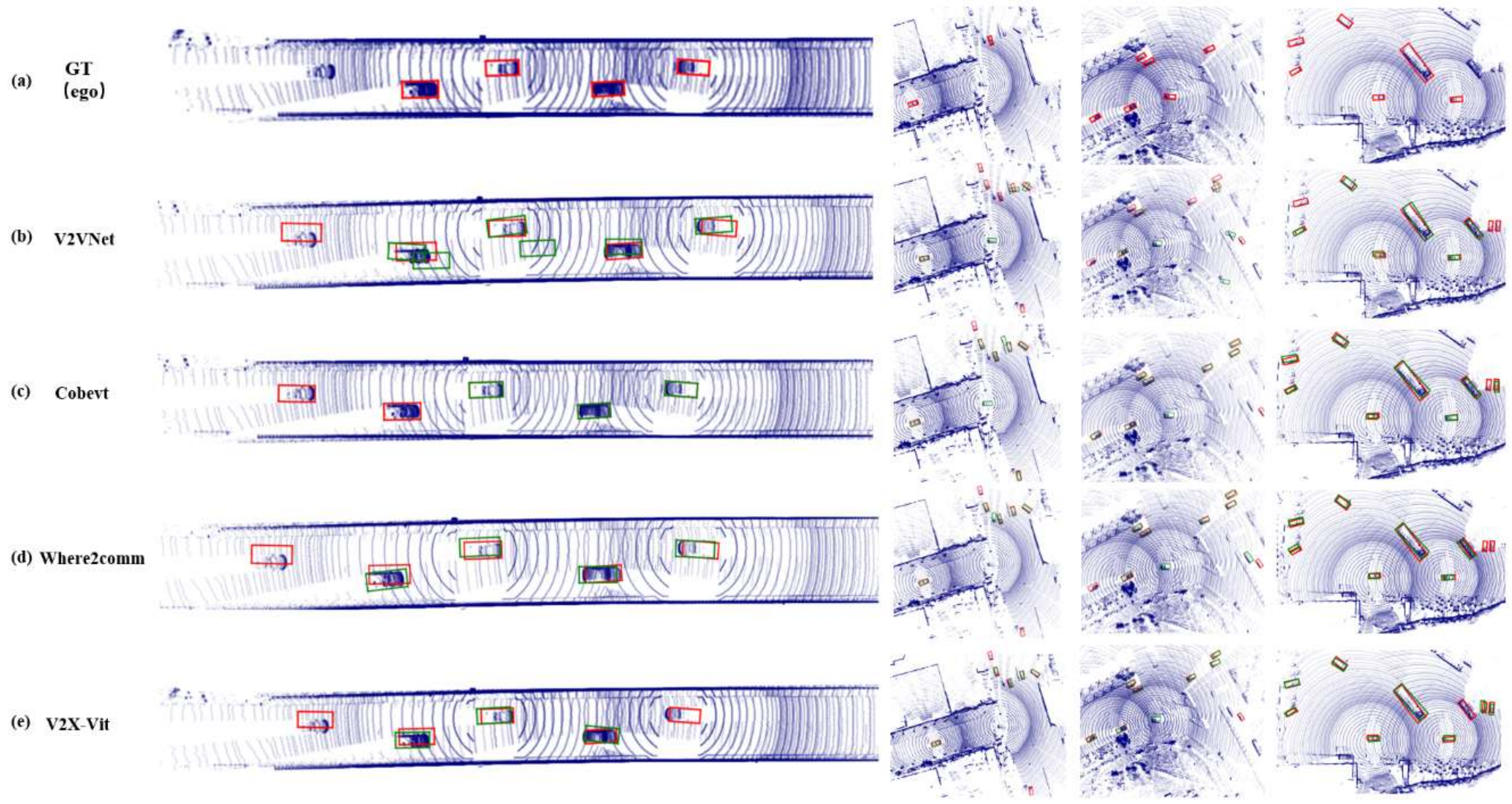
Agent	Sensor	Sensor Model	Detail
2*Vehicle	LiDAR	RoboSense Ruby Plus(*1)	128 beams, 10Hz capture frequency, 360°horizontal FOV, -25°to +15°vertical FOV, < 200m range
	Camera	Sensing Cmaera(*5)	front-wide: SG8S-AR0820C-5300-G2A-Hxxx, 8MP, HFOV30°, front-broad: SG8S-AR0820C-5300-G2A-Hxxx, 8MP, HFOV120°, left&right: SG2-AR0231C-0202-GMSL-Hxxx, 2MP, HFOV100°, back: SG2-AR0233C-5200-G2A-Hxxx, 2mp, HFOV121°
	GPS&IMU	Intelligent Car Built-in GPS System(*1)	100HZ
UAV	LiDAR	RoboSense Helios32(*1)	32 beams, 10Hz capture frequency, 360°horizontal FOV, -55°to +15°vertical FOV,< 150m range
	Camera	USB Camera(*1)	front: RER-USBGS1200P02, 2MP, HFOV120°
	GPS&IMU	DJI M350 RTK Built-in GPS System(*1)	GPS + GLONASS + BeiDou + Galileo, 100HZ

Benchmarks & Results

- Two main benchmarks: **AGC-V2V** (vehicle–vehicle) and **AGC-VUC** (vehicle–UAV).
- Baselines include **PointPillars, V2VNet, Where2Comm, CoBEVT,** and **V2X-ViT**.
- Metrics: mAP@0.5, mAP@0.7, ΔUAV improvement. $\Delta_{\text{UAV}} = \frac{1}{2}\left[\left(m_{0.5}^{V2U} - m_{0.5}^{V2V}\right) + \left(m_{0.7}^{V2U} - m_{0.7}^{V2V}\right)\right]$
- Unified BEV representation ensures fair cross-domain comparison.

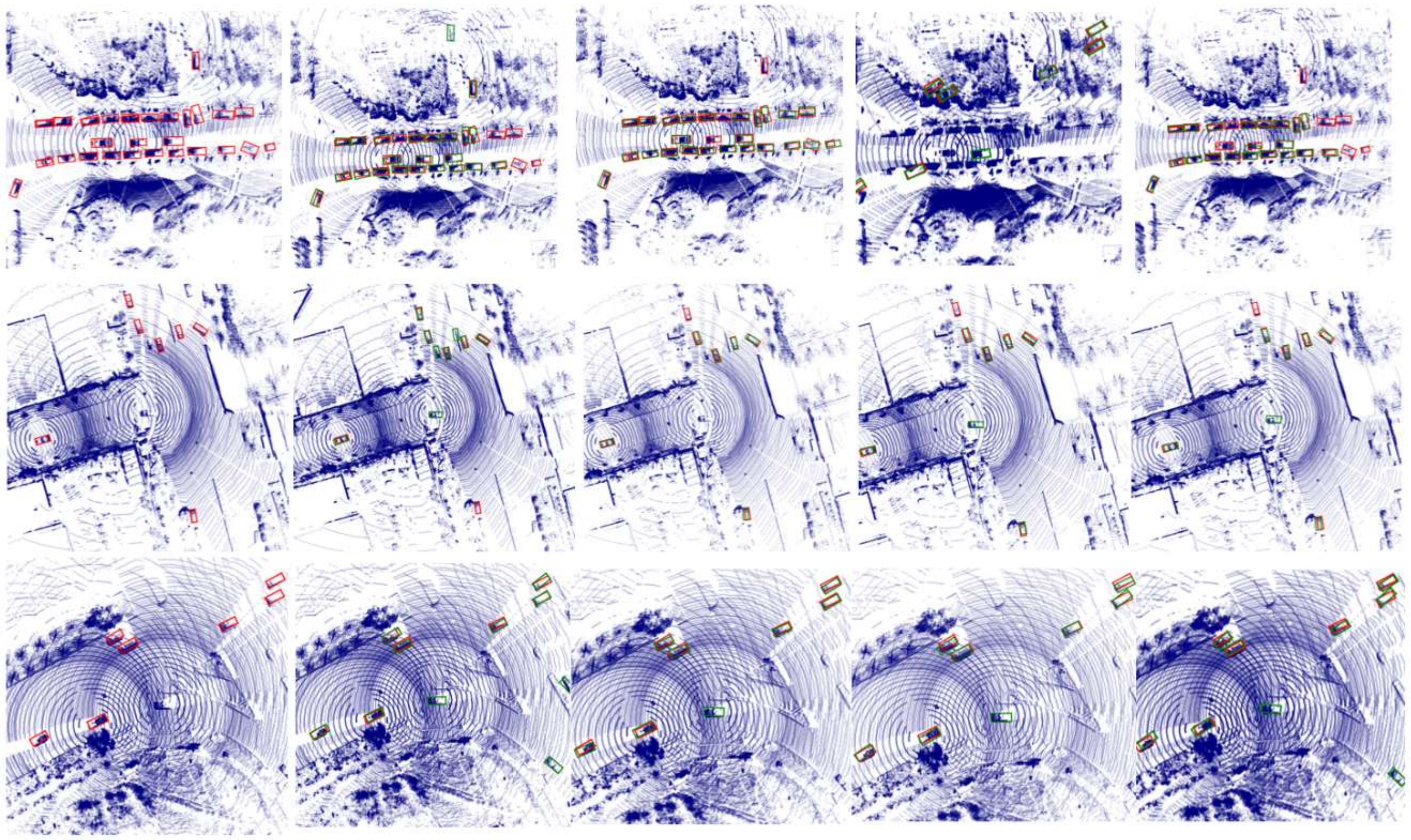
Table 3: 3D Detection Performance (%) on AGC-V2V.

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



Benchmarks & Results

- Introducing UAV data yields **consistent mAP improvements** (+0.6–11.5%).
- **Performance gain** demonstrates complementary aerial-ground perspectives.
- AGC-Drive sets **new standard** for air-ground collaborative perception benchmarks.



(a) GT (b) V2VNet (c) Cobevt (d) Where2comm (e) V2X-Vit

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

Co-Mode	Model	V2V mAP@0.5	V2V mAP@0.7	V2U mAP@0.5	V2U 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

Limitations & Future Work

- UAV LiDAR sparsity limits fine-grained perception of small objects.
- Plan to adopt higher-resolution LiDAR and multi-UAV collaboration.
- Expand dataset to various weather, lighting, and traffic conditions.
- Integrate radar and driver-state data for richer multimodal perception.

uncommon category

diverse application scenarios

diverse weather conditions

corner case