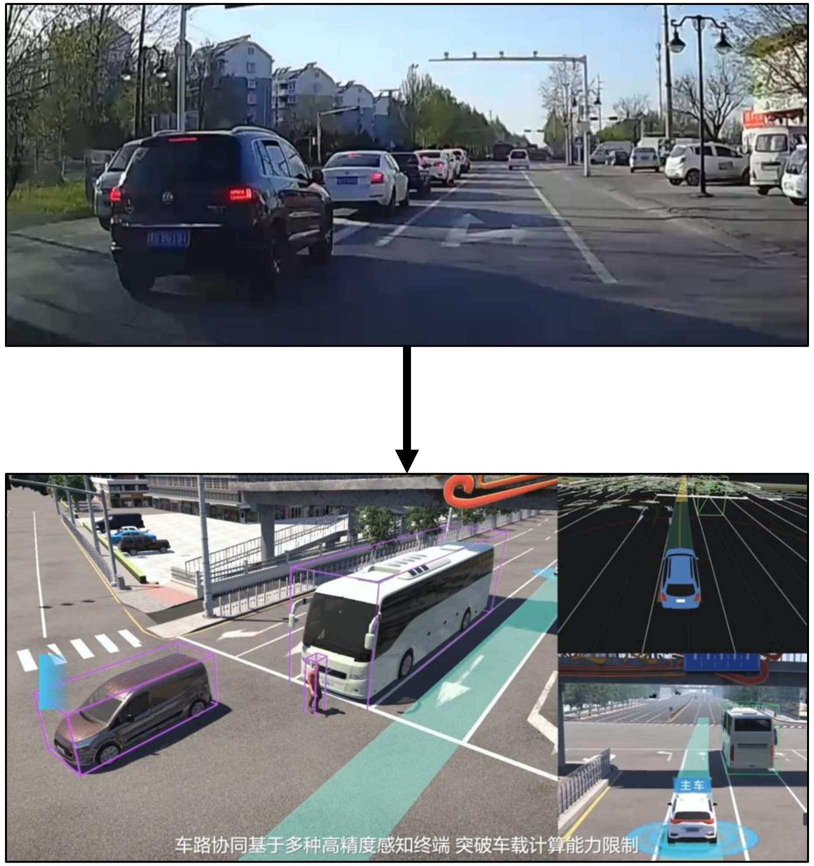
V2X-Radar: A Multi-modal Dataset with 4D Radar for Cooperative Perception

Lei Yang[1,2], Xinyu Zhang[1], Jun Li[1] , Chen Wang[3] , Jiaqi Ma[4] , Zhiying Song[1] , Tong Zhao[1]
Ziying Song[5] , Li Wang[1] , Mo Zhou[1], Yang Shen[1] , Kai Wu[6] ,Chen Lv[2]

[1] School of Vehicle and Mobility, Tsinghua University; [2] Nanyang Technological University; [3] CUMTB
[4] University of California, Los Angeles; [5] Beijing Jiaotong University; [6] ByteDance

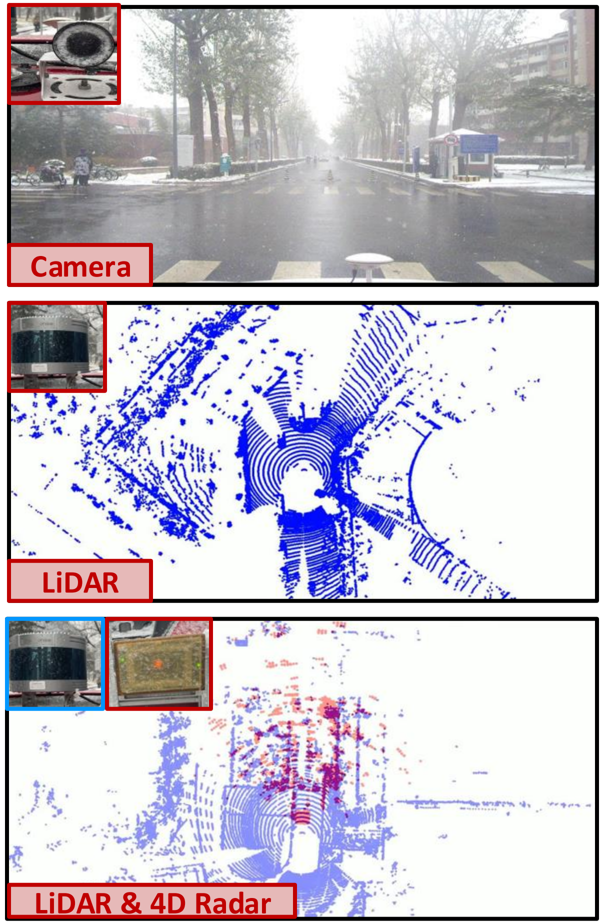
Motivation

- **Cooperative perception** mitigates occlusion and remote sensing, enabling more **safe autonomous driving**.
- Integrating Camera, LiDAR, and **4D Radar** ensures robust perception across all weather and lighting conditions.



Cooperative Perception

Diverse Weather Conditions



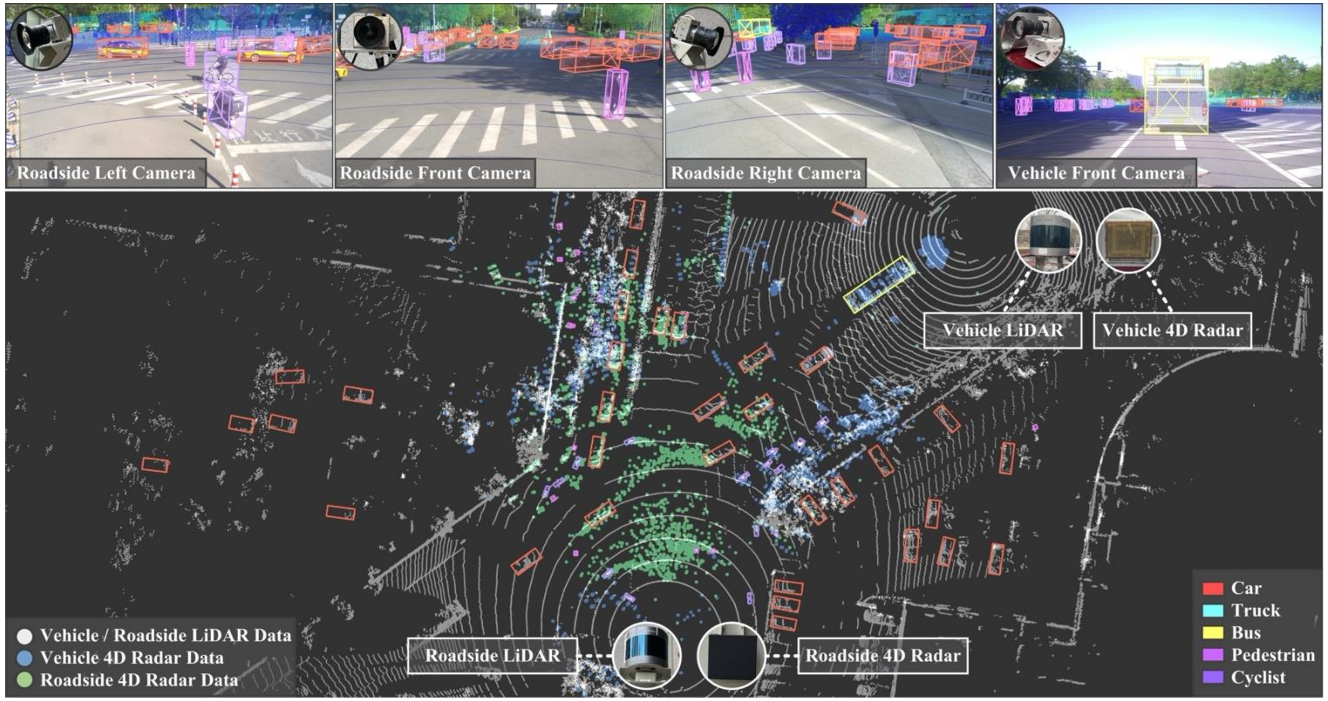
Multi-modal Data

Contributions

- **V2X-Radar** — the first large-scale real-world cooperative dataset integrating **Camera, LiDAR, and 4D Radar**.
- Comprehensive **all-weather** & **all-time** data, collected under diverse conditions and complex traffic scenes.
- Supports **vehicle-side**, **roadside**, and **cooperative** perception tasks, with extensive multi-modal benchmarks.

Comparisons of our proposed V2X-Radar with existing V2X datasets.

Dataset	Location	Sensor	V2X	Real	4D Radar	Adverse Weather	Day&Night	#LiDAR	#Images	#4D Radar	#3D Box
OpenV2V [40]	Carla	L&C	V2V	✗	✗	✗	✗	11K	44K	0	233K
V2X-Set [39]	Carla	L&C	V2V&I	✗	✗	✗	✗	11K	44K	0	233K
DAIR-V2X [51]	China	L&C	V2I	✓	✗	✗	✗	39K	39K	0	464K
V2X-Seq [53]	China	L&C	V2I	✓	✗	✗	✗	39K	39K	0	464K
V2V4Real [38]	USA	L&C	V2V	✓	✗	✗	✗	20K	40K	0	240K
TUMTraf-V2X[59]	DE	L&C	V2I	✓	✗	✗	✓	2K	5K	0	29.38K
RCooper [7]	China	L&C	I2I	✓	✗	✗	✗	30K	50K	0	310K
V2X-Radar	China	L&C&R	V2I	✓	✓	✓	✓	20K	40K	20K	350K



Sample data from the V2X-Radar dataset

V2X-Radar Dataset

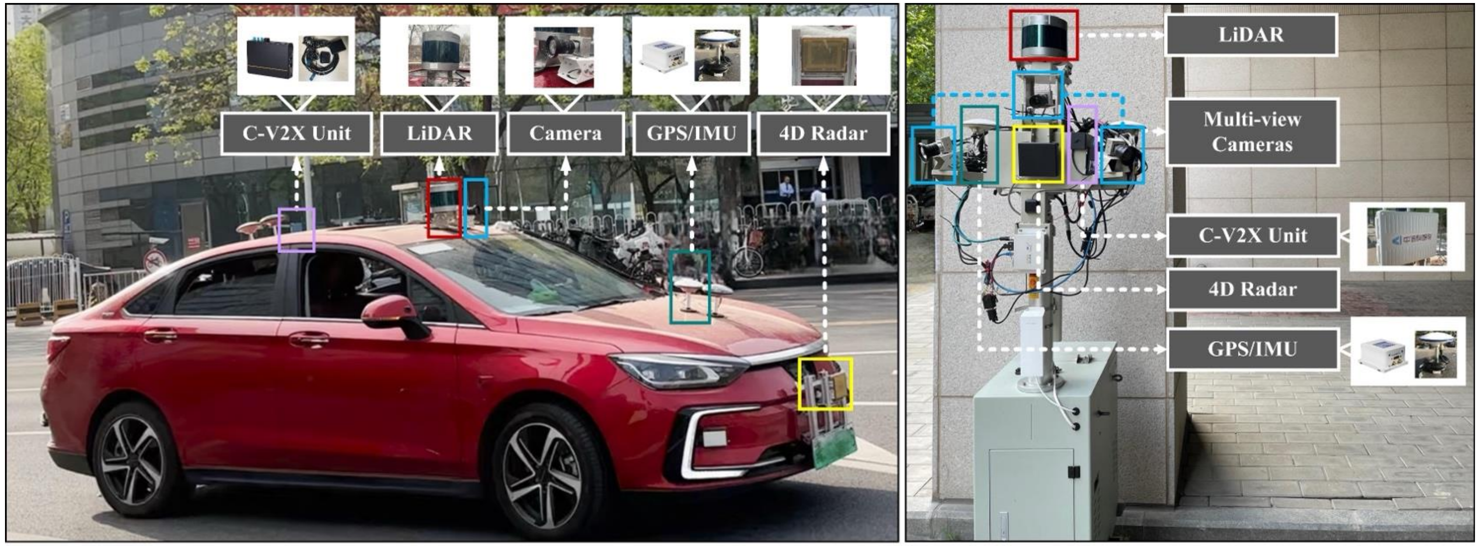
Sensor Setup → Calibration and Registration → Data Collection → Tasks → Benchmarks

Vehicle-side platform and **Intelligent roadside unit**, both equipped with:

- **4D Radar** for velocity-aware sensing
  LiDAR for dense 3D geometry;
- **Multi-view Cameras** for appearance and semantic cues;
- **GPS/IMU** for high-precision localization;
- **C-V2X module** for real-time data transmission;

Sensor specifications in V2X-Radar dataset.

Agent	Sensor	Sensor Model	Details
Infra.	LiDAR	RoboSense RS-Ruby-80 (×1)	80 beams, 360° horizontal FOV, −25° to + 25° vertical FOV
	Camera	Basler acA1920-40gc (×3)	RGB, 1536×864 resolution
	4D Radar	OCULI EAGLE (×1)	79.0GHz, −56° to + 56° horizontal FOV, −22° to + 22° vertical FOV
	C-V2X Unit	VU4004 (×1)	PC5/4G LTE/V2X Protocol
	GPS/IMU	XW-GI5651 (×1)	1000Hz update rate, Double-Precision
Vehicle	LiDAR	RoboSense RS-Ruby-80 (×1)	80 beams, 360° horizontal FOV, −25° to + 25° vertical FOV
	Camera	Basler acA1920-40gc (×1)	RGB, 1920×1080 resolution
	4D Radar	Arbe Phoenix (×1)	77GHz, −50° to + 50° horizontal FOV, −15° to + 15° vertical FOV
	C-V2X Unit	VU4004 (×1)	PC5/4G LTE/V2X Protocol
	GPS/IMU	XW-GI5651 (×1)	1000Hz update rate, Double-Precision

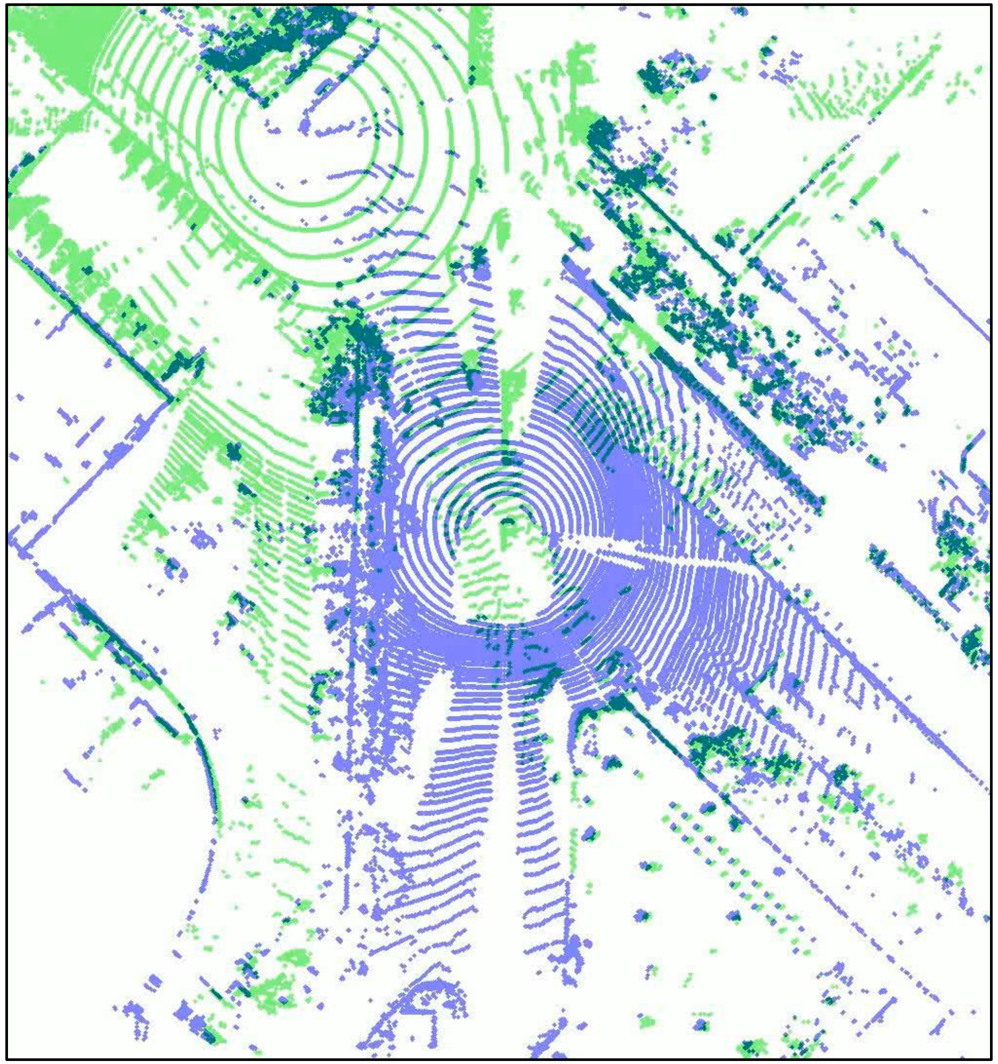


(a) Vehicle-side Platform (b) Intelligent Roadside Unit

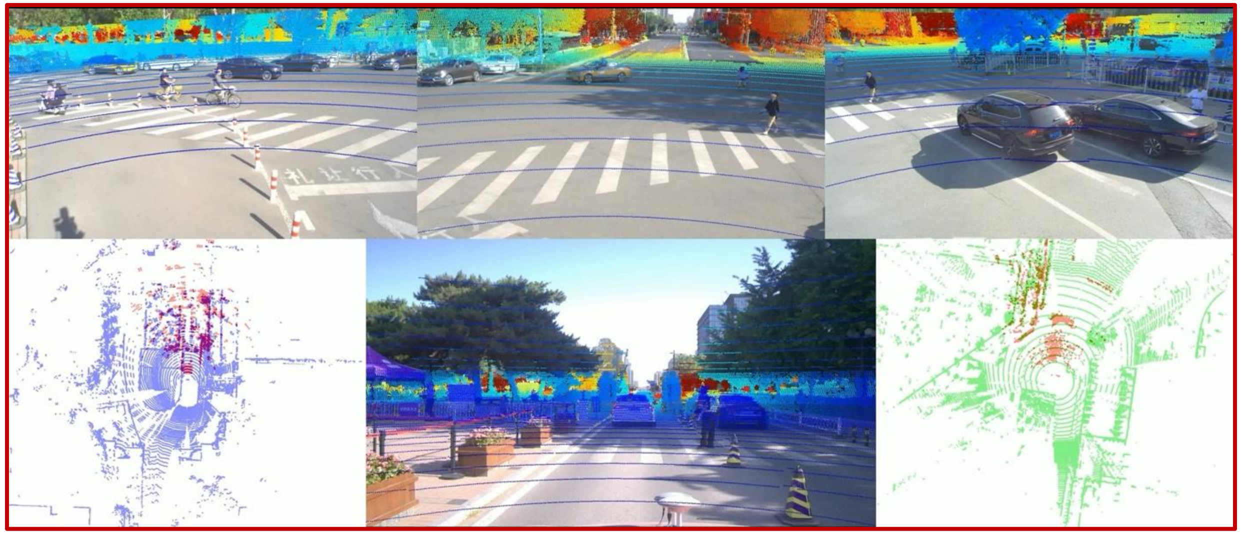
The sensor configuration on the connected vehicle-side platform and the intelligent roadside unit.

V2X-Radar Dataset

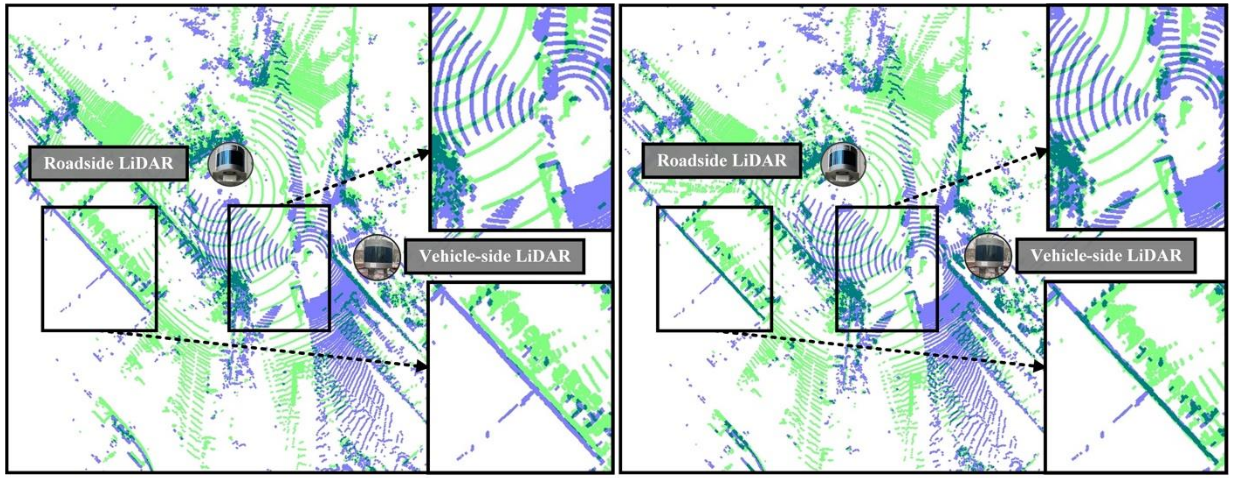
Sensor Setup → **Calibration and Registration** → **Data Collection** → **Tasks** → **Benchmarks**



Visualization of vehicle-infrastructure point cloud registration results.



Visualization of single-agent sensor calibration results



(a) Initial point cloud registration based on RTK

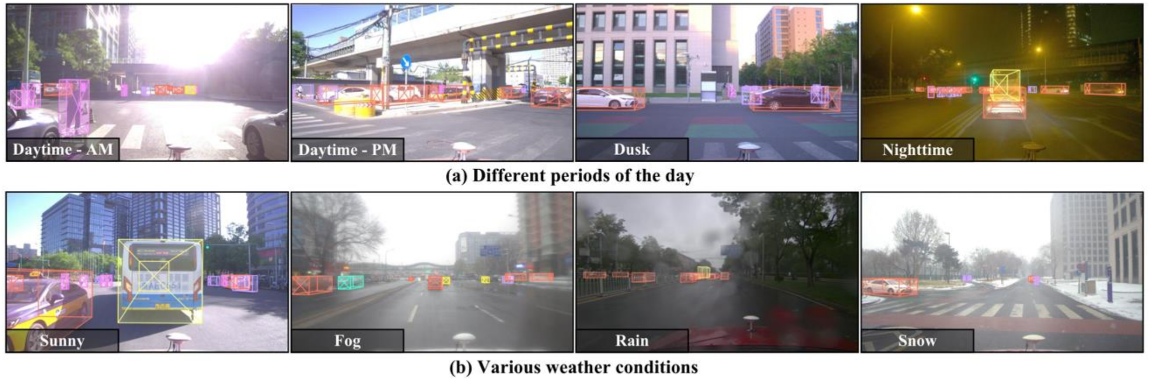
(b) Refined point cloud registration with CBM and manual adjustment

V2X-Radar Dataset

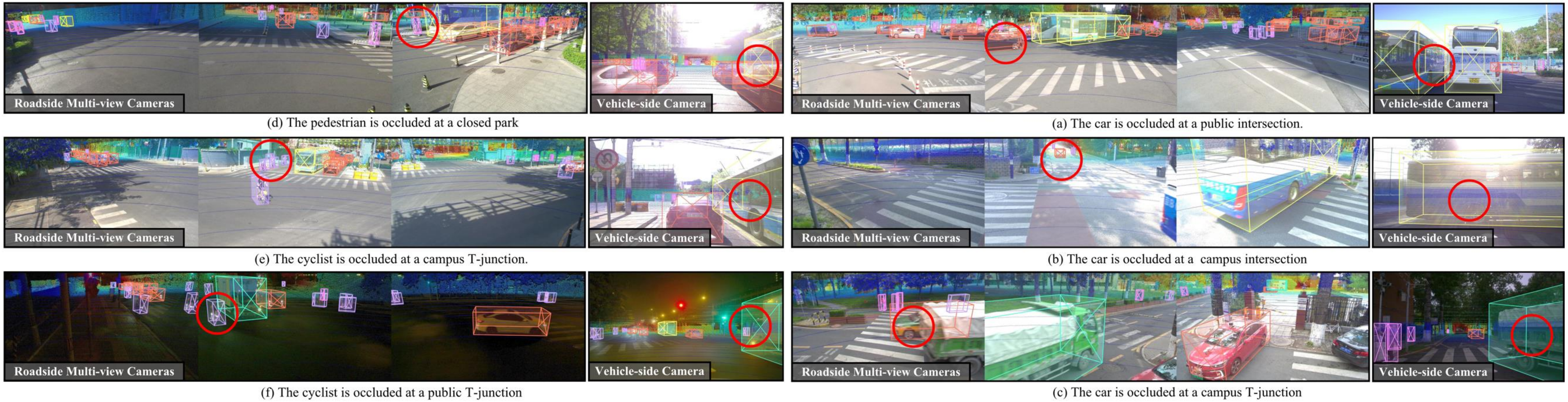
NEURAL INFORMATION PROCESSING SYSTEMS

Sensor Setup → Calibration and Registration → **Data Collection** → Tasks → Benchmarks

- Spans 9 months, with 15 hours and 540K frames;
- **Various Time-of-Day:** morning, afternoon, dusk, night;
- **Diverse Weather Conditions:** sunny, foggy, rainy, and snowy;
- **Rich Corner Cases** : vehicle and pedestrian occlusion, complex intersections, and multi-agent interactions.



Data collection across different periods and various weather conditions.



Diverse corner cases for single-vehicle autonomous driving.

V2X-Radar Dataset

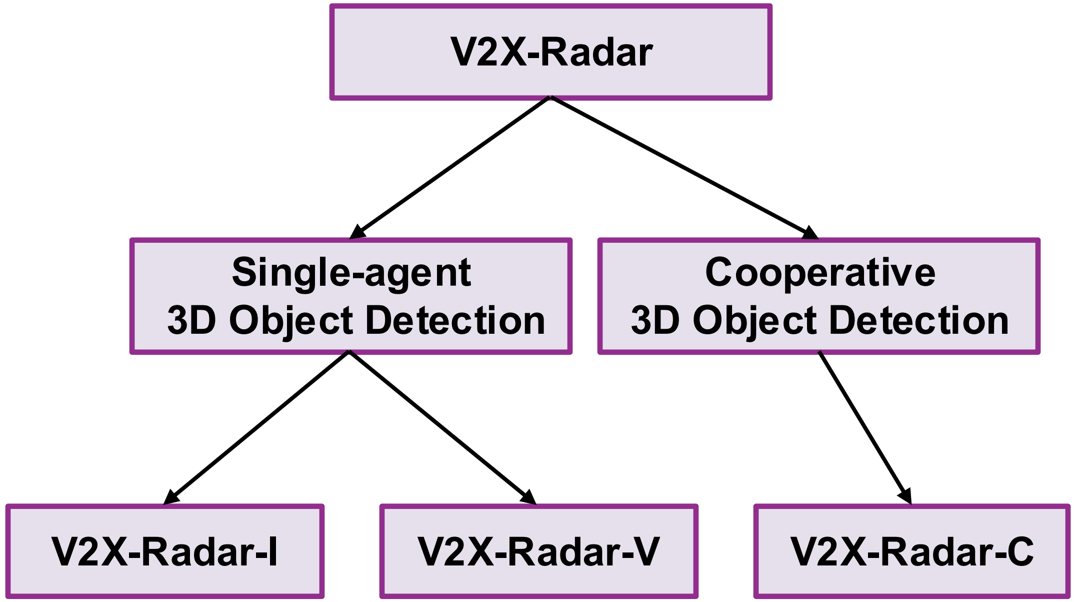
Sensor Setup → Calibration and Registration → Data Collection → **Tasks** → Benchmarks

☐ Single-agent 3D Object Detection

- Conducted on vehicle-side (V2X-Radar-V) and roadside (V2X-Radar-I) subsets;
- **Evaluation metrics**: AP@IoU=0.5 / 0.7 across vehicles, pedestrians, and cyclists;
- **Challenges**: Single-modal Encoding, Multi-modal Fusion.

☐ Cooperative 3D Object Detection

- Conducted on V2X-Radar-C subset, fusing vehicle and roadside data.
- **Evaluation metrics**: AP@IoU=0.5 / 0.7 under both temporal synchrony and asynchrony (different transmission delay);
- **Challenges**: Spatial Asynchrony, Temporal Asynchrony



dataset composition of V2X-Radar dataset

V2X-Radar Dataset

Sensor Setup → Calibration and Registration → Data Collection → Tasks → **Benchmarks**

(1) Single-agent 3D Object Detection

- Roadside Perception on V2X-Radar-I
- Single-vehicle Perception on V2X-Radar-V

LiDAR > 4D Radar > Camera

Roadside > Single Vehicle

Roadside 3D object detection benchmarks on V2X-Radar-I under homogeneous split.

Method	M	Vehicle (IoU = 0.7/0.5) ↑			Pedestrian (IoU = 0.5/0.25) ↑			Cyclist (IoU = 0.5/0.25) ↑		
		Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
Pointpillars [11]	L	78.00/88.69	71.24/81.16	71.24/81.16	53.87/69.61	52.94/67.09	52.94/67.09	80.01/87.48	72.67/78.60	72.67/78.60
SECOND [41]	L	81.61/91.06	74.34/83.47	74.34/83.47	57.56/74.64	55.39/72.27	55.39/72.27	80.53/87.68	74.06/80.62	74.06/80.62
CenterPoint [50]	L	86.44/94.04	78.98/84.26	78.98/84.26	67.90/84.74	65.39/81.35	65.39/81.35	90.26/92.91	82.87/85.51	82.87/85.51
PV-RCNN [21]	L	88.83/94.11	81.39/86.61	81.39/86.61	77.13/86.39	74.66/83.87	74.66/83.87	91.82/94.44	84.47/87.08	84.46/87.08
SQDNet [55]	L	**89.12/95.10**	**81.48/86.94**	**81.48/86.94**	**78.02/88.47**	**74.79/84.19**	**74.79/84.19**	**92.13/95.43**	**84.53/87.59**	**84.53/87.59**
Fade3D [48]	L	81.03/90.43	73.56/81.72	73.56/81.72	66.07/82.85	64.55/79.07	64.55/79.07	88.43/90.76	75.81/82.17	75.81/82.17
SMOKE [16]	C	22.05/58.43	20.72/56.36	20.69/56.31	8.26/25.64	7.68/24.36	7.63/24.30	12.50/38.29	11.28/36.40	11.24/36.36
BEVDepth [12]	C	45.01/69.25	42.23/66.81	42.21/66.75	30.64/61.46	29.13/59.11	29.10/59.05	39.85/68.71	38.52/67.05	38.43/67.02
BEVHeight [43]	C	47.91/72.45	45.53/69.48	45.49/67.44	32.08/64.06	29.78/59.79	29.68/59.74	42.97/71.63	41.34/69.11	41.30/69.07
BEVHeight++[42]	C	48.48/73.81	47.92/70.36	47.88/70.32	33.05/66.12	32.30/64.67	32.32/64.66	45.19/74.52	44.20/72.04	44.14/72.01
RDIoU [20]	R	61.38/80.03	54.89/72.72	54.89/72.72	43.82/72.03	42.11/69.65	42.11/69.65	40.74/67.31	36.68/60.83	36.68/60.83
RPFA-Net [34]	R	64.79/82.58	58.01/75.36	58.01/75.36	51.64/78.05	49.51/73.91	49.51/73.91	45.86/71.81	41.66/64.95	41.66/64.95

Single-vehicle 3D object detection benchmarks on V2X-Radar-V under homogeneous split.

Method	M	Vehicle (IoU = 0.7/0.5) ↑			Pedestrian (IoU = 0.5/0.25) ↑			Cyclist (IoU = 0.5/0.25) ↑		
		Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
Pointpillars [11]	L	75.66/83.52	68.80/77.07	68.80/77.07	41.89/46.34	38.16/43.18	38.16/43.18	78.63 /83.14	65.24/69.77	65.24/69.77
SECOND [41]	L	78.35/84.34	71.31/79.75	71.31/79.75	43.74/49.75	39.07/45.91	39.07/45.91	82.88/85.12	68.27/73.22	68.27/73.22
CenterPoint [50]	L	80.87/87.74	72.19/81.42	72.19/81.42	55.27/62.63	50.59/58.81	50.59/58.81	88.21/92.48	75.26/79.44	75.26/79.44
PV-RCNN [21]	L	88.27/89.12	79.38/84.31	79.38/84.31	67.04/69.81	**58.83/65.78**	**58.83/65.78**	**89.48/93.49**	78.01/81.07	78.01/81.07
SQDNet [55]	L	**89.02/90.65**	**79.65/85.10**	**79.65/85.10**	**68.23/72.85**	58.79/65.55	58.79/65.55	88.04/91.85	**79.46/82.95**	**79.46/82.95**
Fade3D [48]	L	79.24/85.77	70.64/78.95	70.64/78.95	57.82/65.92	51.88/60.94	51.88/60.94	84.95/88.29	69.93/75.80	69.93/75.80
SMOKE [16]	C	9.86/31.92	8.61/26.41	8.28/24.36	0.23/2.02	0.29 /1.89	0.29/1.89	0.39/7.23	0.37/4.96	0.37/4.13
BEVDepth [12]	C	16.91/41.63	15.47/39.68	15.02/37.83	9.92/29.98	8.51/27.76	8.49/27.72	12.18/47.20	9.46/39.34	9.30/39.15
BEVHeight [43]	C	16.58/40.32	15.32/39.15	14.08/37.30	9.49/28.50	8.48/26.46	8.39/26.57	9.58/44.56	7.35/35.04	7.26/34.91
BEVHeight++[42]	C	17.47/43.68	15.53/42.24	14.77/41.58	10.43/ 31.15	9.36/28.85	9.32/28.73	12.99/49.08	9.91/41.10	9.83/40.94
RDIoU [20]	R	41.11/72.67	29.03/54.27	28.37/52.02	10.72/28.59	9.97/26.88	9.84/26.78	14.74/44.57	10.81/31.15	10.67/30.91
RPFA-Net [34]	R	42.77/75.79	30.44/57.27	29.34/55.06	11.51/30.54	10.37/28.15	10.29/27.43	17.03/46.31	11.98/33.96	11.91/33.77

V2X-Radar Dataset

NEURAL INFORMATION PROCESSING SYSTEMS

Sensor Setup → Calibration and Registration → Data Collection → Tasks → **Benchmarks**

(2) Cooperative 3D Object Detection

- ☐ Temporal Synchrony Evaluation

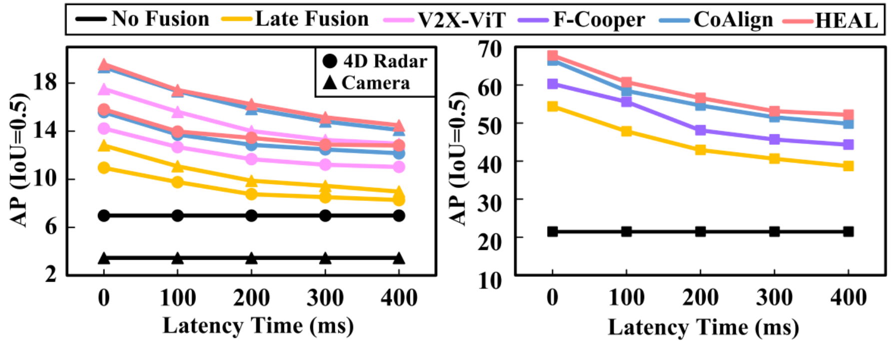
→ Cooperative perception without transmission delay (Sync)

- ☐ Temporal Asynchrony Evaluation

→ Simulated 100, 200, 300, 400 ms latency between vehicle and roadside (Async)

Cooperative 3D object detection benchmarks for vehicle category on V2X-Radar-C.

Method	M	Sync. (AP@IoU = 0.7 / 0.5) ↑				Async. (AP@IoU = 0.7 / 0.5) ↑				AM
		Overall	0-30m	30-50m	50-100m	Overall	0-30m	30-50m	50-100m	(MB)
No Fusion	L	17.40 / 21.44	18.54 / 22.73	18.60 / 22.22	13.64 / 18.15	17.40 / 21.44	18.54 / 22.73	18.60 / 22.22	13.64 / 18.15	0
Late Fusion	L	35.03 / 54.38	52.93 / 67.09	35.46 / 49.56	20.51 / 32.58	22.11 / 47.84	33.13 / 59.97	20.84 / 42.05	14.78 / 28.27	0.003
Early Fusion	L	49.04 / 58.89	63.74 / 74.13	43.64 / 52.73	28.71 / 34.38	27.22 / 51.09	35.67 / 64.38	22.05 / 43.26	19.08 / 29.93	1.25
F-Cooper [2]	L	45.33 / 60.27	61.87 / 74.91	38.95 / 56.64	21.32 / 35.85	32.03 / 55.56	44.81 / 70.11	23.64 / 50.01	17.31 / 32.48	0.25
CoAlign [18]	L	55.56 / 66.42	63.27 / 70.81	54.15 / 66.75	43.01 / 58.36	35.07 / 58.43	39.61 / 63.32	31.83 / 56.64	31.00 / 50.65	0.25
HEAL [17]	L	**57.78 / 67.75**	**65.59 / 72.78**	**56.23 / 69.24**	**45.32 / 57.79**	**39.27 / 60.75**	**43.95 / 65.72**	**36.04 / 59.99**	**34.52 / 51.41**	**0.25**
No Fusion	C	2.61 / 3.45	1.56 / 5.08	0.29 / 0.69	0.10 / 0.22	2.61 / 3.45	1.56 / 5.08	0.29 / 0.69	0.10 / 0.22	0
Late Fusion	C	7.56 / 12.82	16.53 / 27.33	1.93 / 2.38	0.92 / 1.87	4.68 / 11.08	10.19 / 23.88	1.14 / 1.94	0.58 / 1.59	0.003
V2X-ViT [39]	C	10.54 / 17.51	19.93 / 32.28	2.64 / 3.42	1.39 / 2.34	7.24 / 15.61	13.58 / 29.19	1.71 / 2.95	0.92 / 2.05	0.25
CoAlign [18]	C	12.67 / 19.33	21.05 / 37.01	2.73 / 3.89	1.41 / 2.45	8.61 / 17.33	14.11 / 33.41	1.74 / 3.27	0.97 / 2.17	0.25
HEAL [17]	C	13.56 / 19.58	24.84 / 34.74	2.88 / 4.02	1.53 / 2.80	9.53 / 17.42	17.43 / 31.23	1.96 / 3.43	0.98 / 2.47	0.25
No Fusion	R	3.06 / 6.98	6.62 / 11.63	0.52 / 1.08	0.26 / 0.87	3.06 / 6.98	6.62 / 11.63	0.52 / 1.08	0.26 / 0.87	0
Late Fusion	R	6.87 / 10.95	11.69 / 20.59	1.15 / 1.93	0.54 / 1.21	4.67 / 9.76	7.83 / 18.39	0.74 / 1.44	0.41 / 1.06	0.003
Early Fusion	R	7.37 / 13.65	13.89 / 24.07	1.49 / 2.19	0.71 / 1.25	5.24 / 12.09	9.76 / 21.51	1.01 / 1.67	0.46 / 1.09	0.32
V2X-ViT [39]	R	8.74 / 14.23	14.92 / 25.23	1.77 / 2.84	0.93 / 1.66	6.01 / 12.68	10.17 / 22.82	1.15 / 2.25	0.61 / 1.46	0.25
CoAlign [18]	R	9.76 / 15.59	17.32 / 26.76	1.86 / 3.08	1.23 / 2.05	6.52 / 13.71	11.49 / 23.63	1.16 / 2.41	0.82 / 1.77	0.25
HEAL [17]	R	10.08 / 15.82	18.24 / 27.02	1.91 / 3.56	1.12 / 1.83	7.56 / 13.97	12.97 / 24.11	1.23 / 2.85	0.75 / 1.62	0.25



(a) async. benchmarks on camera & 4D Radar (b) async. benchmarks on LiDAR

Cooperative 3D object detection benchmarks on V2X-Radar-C under different transmission delay.

V2X-Radar Dataset

Sensor Setup → Calibration and Registration → Data Collection → Tasks → **Benchmarks**

(3) Adverse-weather Robustness through 4D Radar

(Evaluation on an adverse-weather subset of vehicle-side data)

- 4D Radar's role in Adverse-weather conditions
- Doppler's role in 4D Radar Perception

Ablation results under adverse weather conditions (rain, heavy fog, and snow)

Method	Modality	Vehicle (IoU = 0.5) ↑			Pedestrian (IoU = 0.25) ↑			Cyclist (IoU = 0.25) ↑		
		Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
Pointpillars [11]	LiDAR	47.23	43.12	42.15	21.12	17.52	16.87	29.32	25.34	24.81
	4D Radar	48.35	44.80	43.91	20.14	16.54	16.20	27.85	23.66	22.68
M2-Fusion [50]	LiDAR + 4D Radar	**53.61**	**50.18**	**49.72**	**25.11**	**20.53**	**19.59**	**33.42**	**28.24**	**27.12**

Ablation results on Doppler's contribution to 4D Radar perception under adverse weather (rain, fog, and snow).

Modality	**Veh. (IoU = 0.5)↑**			**Ped. (IoU = 0.25)↑**			**Cyc. (IoU = 0.25)↑**		
	Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
LiDAR	47.23	43.12	42.15	**21.12**	**17.52**	**16.87**	**29.32**	**25.34**	**24.81**
4D Radar w/o Doppler	45.22	41.39	40.92	19.98	14.30	14.24	24.68	21.52	20.56
4D Radar w/ Doppler	**48.35**	**44.80**	**43.91**	20.14	16.54	16.20	27.85	23.66	22.68

Vehicle-side data samples under adverse weather conditions

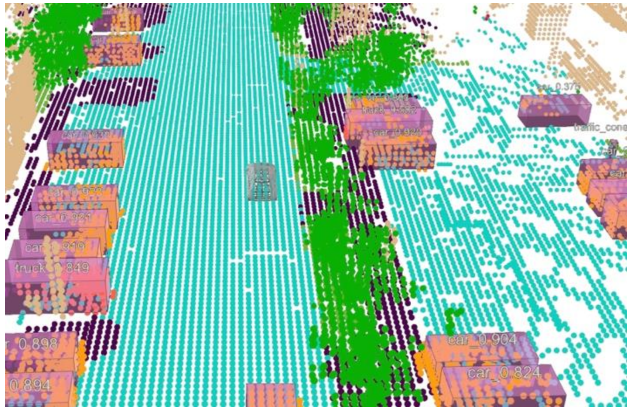
Conclusion

- We introduced **V2X-Radar**, the first real-world multi-modal cooperative dataset featuring **4D Radar**.

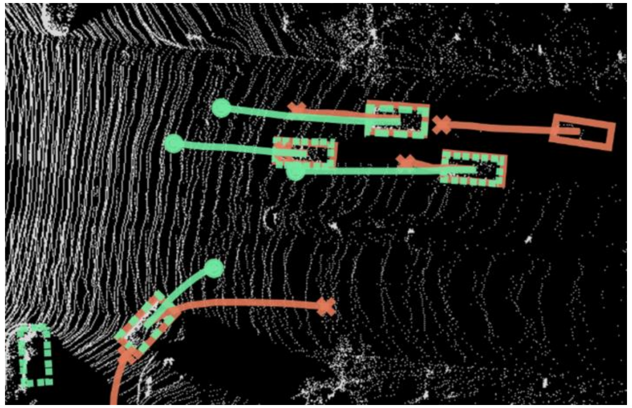
- The dataset reveals **two key insights**:
  - **Delay Sensitivity:** Current cooperative methods degrade under asynchronous communication.
  - **4D Radar Effectiveness:** Provides stable perception in adverse weather with 4D Radar.

- V2X-Radar bridges the **4D Radar** gap for cooperative perception and serves as a platform for **robust perception research**.

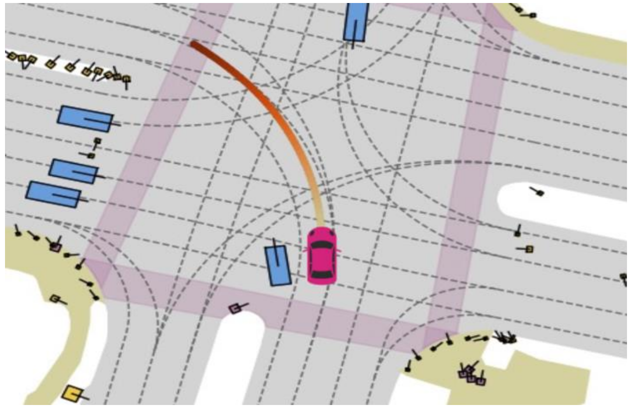
Future Works



Cooperative Occupancy Perception



Cooperative Trajectory Prediction



E2E Cooperative Driving

Resources

NEURAL INFORMATION PROCESSING SYSTEMS

Full Paper: https://arxiv.org/abs/2411.10962

CodeBase: https://github.com/yanglei18/V2X-Radar

Dataset: https://huggingface.co/datasets/yanglei18/V2X-Radar

Full Paper

Code Base

Dataset

WeChat

Thank you !