UrbanIng-V2X: A Large-Scale Multi-Vehicle, Multi-Infrastructure Dataset Across Multiple Intersections for Cooperative Perception

Karthikeyan Chandra Sekaran[1]†, Markus Geisler[1]†, Dominik Rößle[1]†,

Adithya Mohan[1], Daniel Cremers[2], Wolfgang Utschick[2],

Michael Botsch[1], Werner Huber[1], Torsten Schön[1]

† Equal contribution. Authors listed in alphabetical order.

[1] Technische Hochschule Ingolstadt

[2] Technical University of Munich

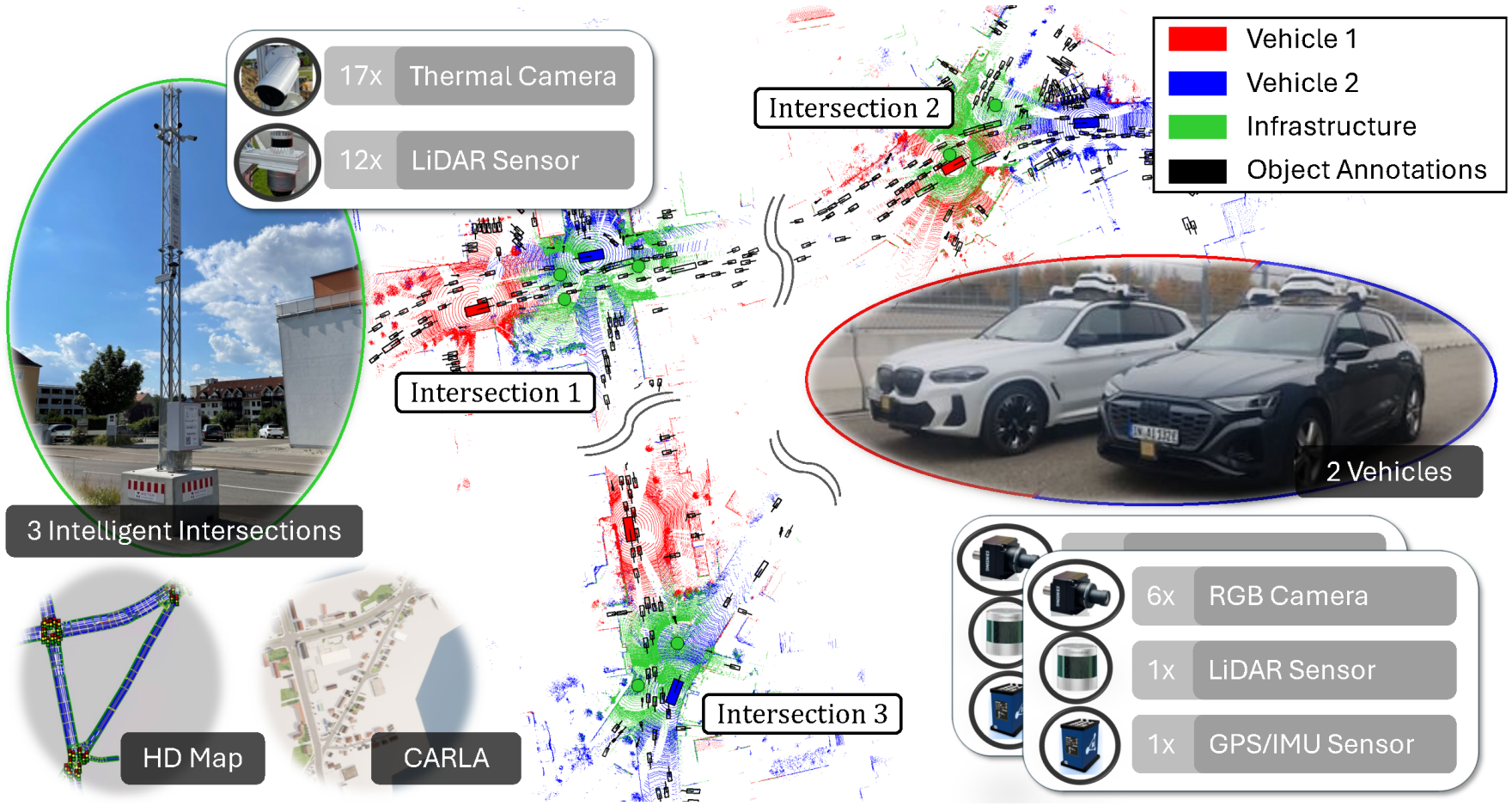
Technische Hochschule Ingolstadt

AImotion Bavaria

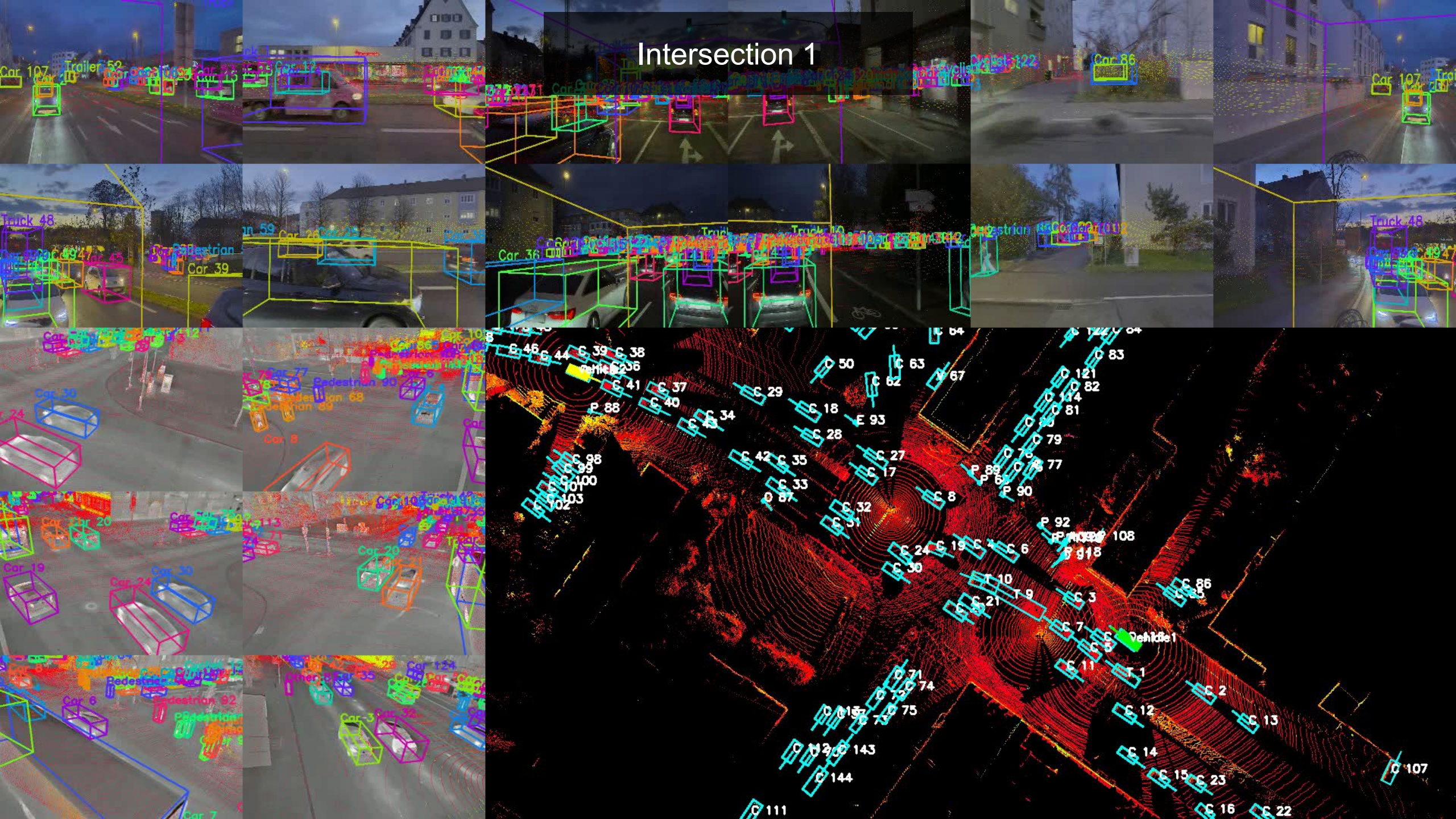
CARISSMA Institute of Automated Driving

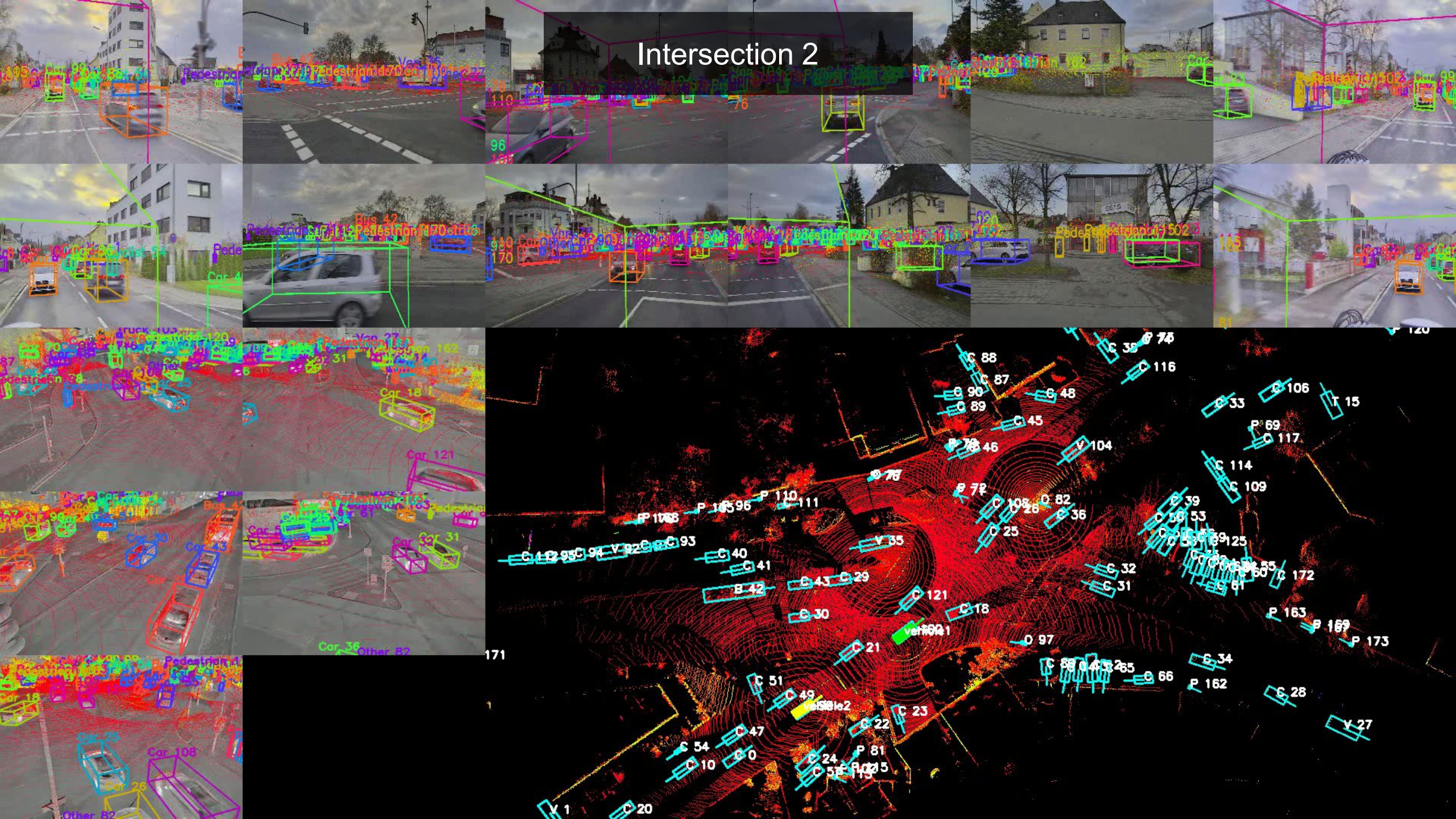
UrbanIng-V2X

A Large-Scale Multi-Vehicle, Multi-Infrastructure Dataset Across Multiple Intersections for Cooperative Perception

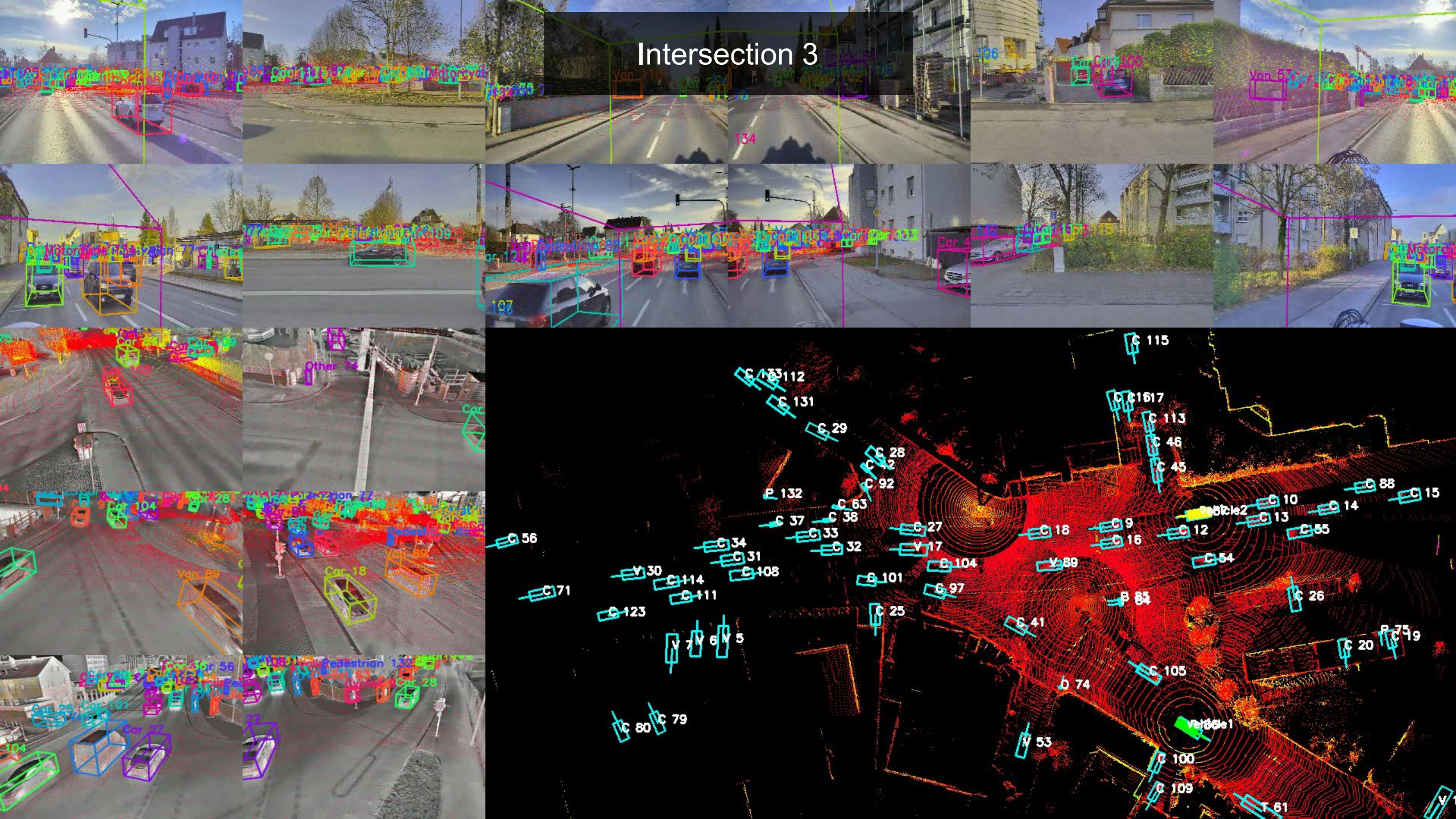


Property	UrbanIng-V2X
V2X	V2V&I
Intersections	3
Vehicles	2
RGB Images	81.6k
IR Images	38.8k
LiDAR Frames	27.2k
3D Boxes	712k
Classes	13
Digital Twin	Yes
Av. Worldwide	Yes
HD Maps	Yes
Attributes	Yes
Track Ids	Yes
Traffic Light	Yes
Sensors (I | V)	10 | 16
City	Ingolstadt
Country	Germany


Intersection 1

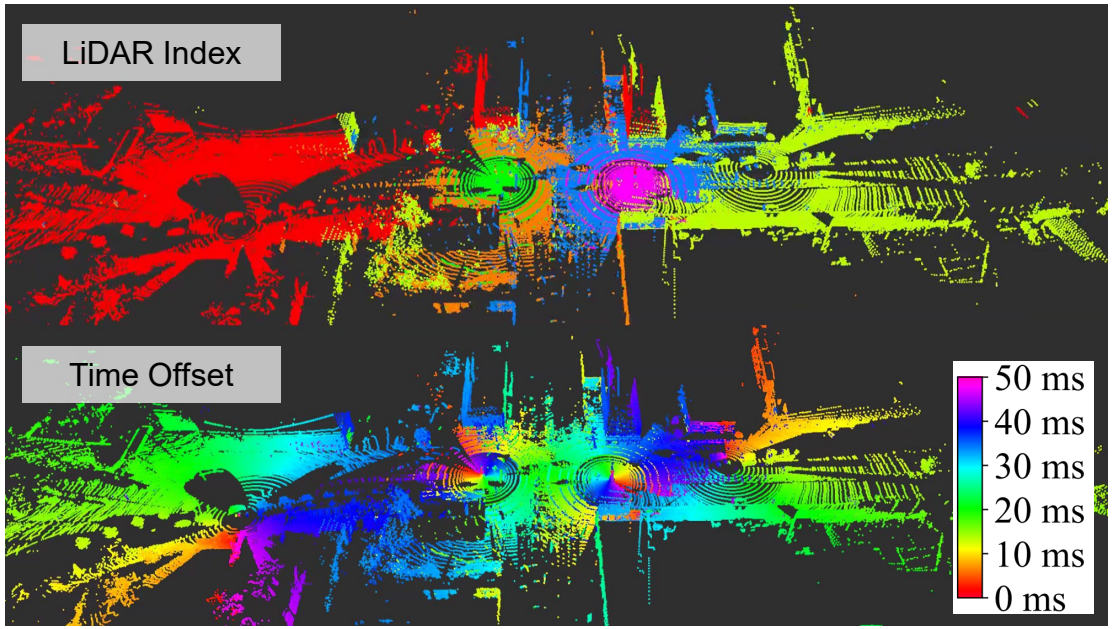

Intersection 2

Intersection 3

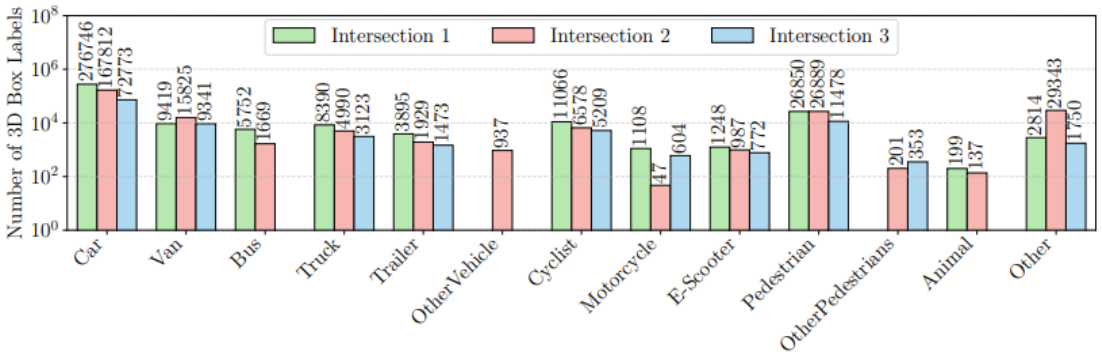


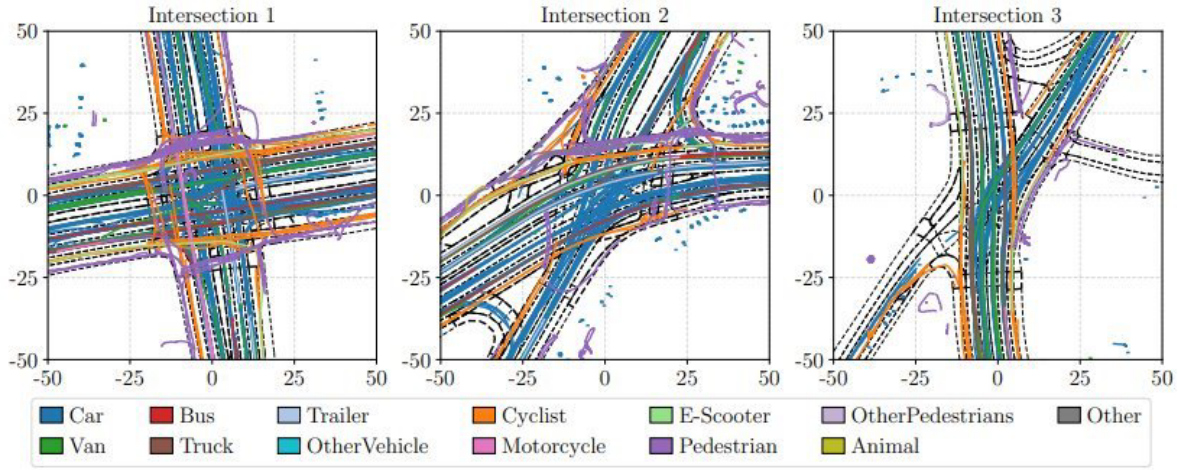
Dataset insights

Space time synchronization; Dataset highlights



Space Time Synchronization

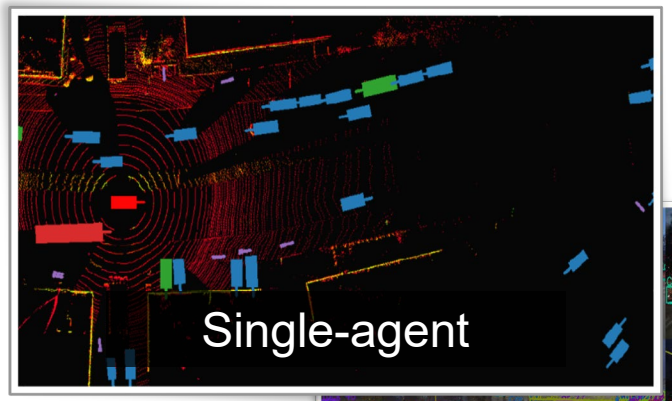


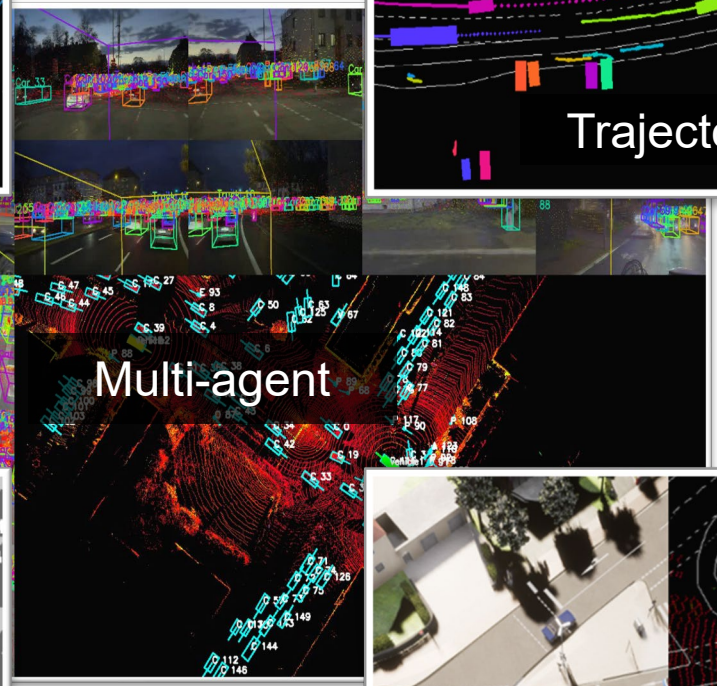


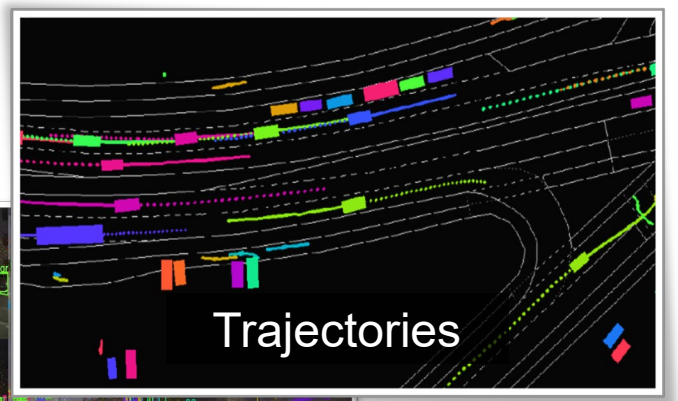
Dataset Statistics

Benchmark tasks

Possible tasks and setups







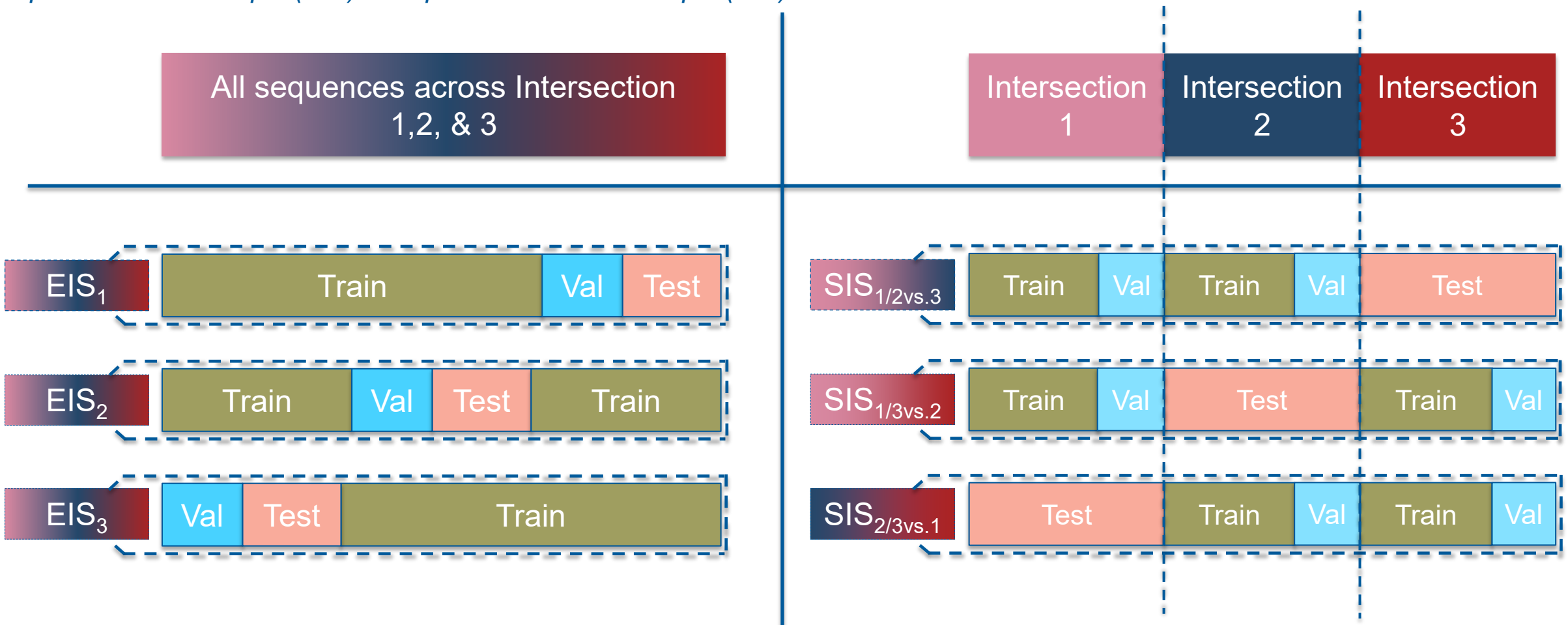
- 3D Multi Object Detection and Tracking
- Trajectory Prediction
- BEV Prediction
- Sim-to-Real Tasks
- HD-Map Tasks

- Single/ Multi Sensor
- Single/ Multi Agent
- With / Without/ Only Infrastructure

… across 3 intersections in one unified dataset

Benchmark results

Equal Intersection Split (EIS) & Separate Intersection Split (SIS)



Non-overlapping validation and test data.
21 training sequences, 6 validation sequences and 7 test sequences.

Training and validation data are from two intersections.
Test split contains the remaining full intersection data.

Benchmark results

Cooperative perception evaluation

Method	AP_{Veh}		AP_{HVeh}		AP_{Ped}		AP_{TWheel}		mAP	
IoU	0.3	0.5	0.3	0.5	0.3	0.5	0.3	0.5	0.3	0.5
No Fusion	49.1	40.9	19.2	17.6	2.0	0.7	18.0	13.8	22.1	18.3
Early Fusion	46.1	41.1	26.8	24.8	6.0	3.5	24.1	21.6	25.8	22.8
Late Fusion	28.7	24.6	9.8	6.9	1.9	0.8	16.7	12.1	14.3	11.1
F-Cooper [1]	52.6	46.7	33.1	24.0	4.6	3.1	25.1	23.2	28.9	24.2
AttFuse [2]	52.7	47.6	34.1	27.8	7.1	4.6	23.7	22.1	29.4	25.5
V2X-ViT [3]	52.0	46.2	32.5	22.2	5.8	3.5	19.7	18.0	27.5	22.5
Where2Comm [4]	50.4	45.8	28.4	25.3	5.1	3.1	23.2	20.9	26.7	23.8
CoBEVT [5]	53.2	46.0	33.8	29.6	5.7	3.3	22.5	20.5	28.8	24.9

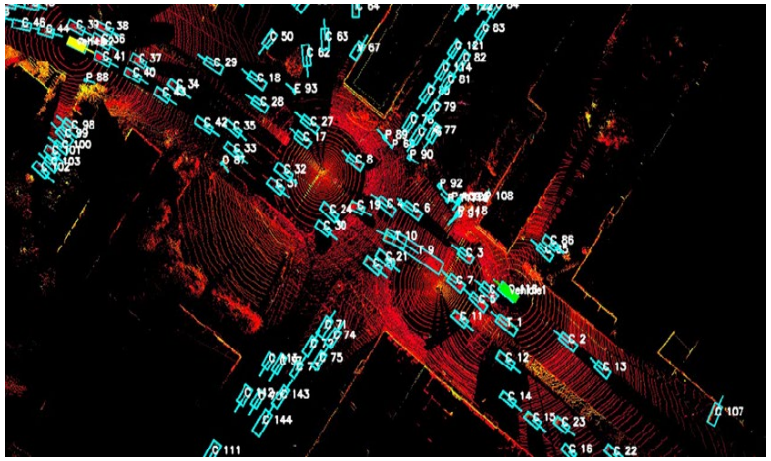
All models were trained and evaluated on the $SIS_{1/2vs.3}$ split.

Data split	AP_{Veh}		AP_{HVeh}		AP_{Ped}		AP_{TWheel}		mAP	
IoU	0.3	0.5	0.3	0.5	0.3	0.5	0.3	0.5	0.3	0.5
EIS_{avg}	74.6	68.7	44.7	37.3	21.8	13.1	38.7	33.0	45.0	38.2
$SIS_{1/2vs.3}$	53.2	46.0	33.8	29.6	5.7	3.3	22.6	20.5	28.8	24.6
$SIS_{1/3vs.2}$	45.1	40.2	14.9	11.3	10.2	6.0	22.0	18.7	23.0	19.1
$SIS_{2/3vs.1}$	64.8	59.1	41.5	31.1	10.8	7.4	22.6	18.2	34.9	28.9

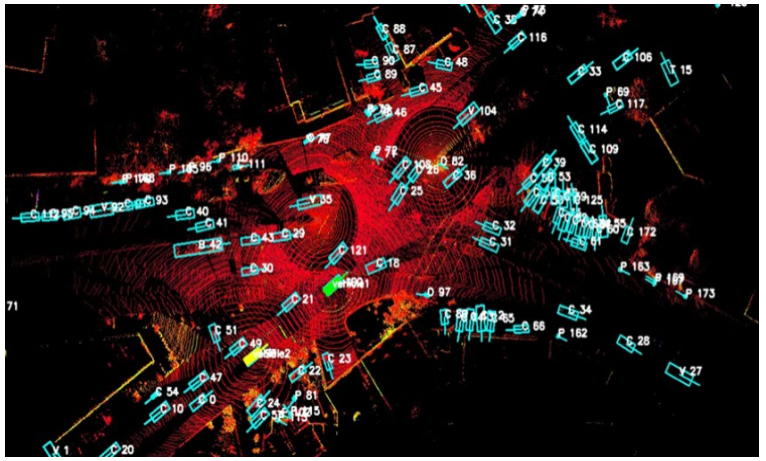
All data splits trained and evaluated on CoBEVT [5].

Check out our dataset

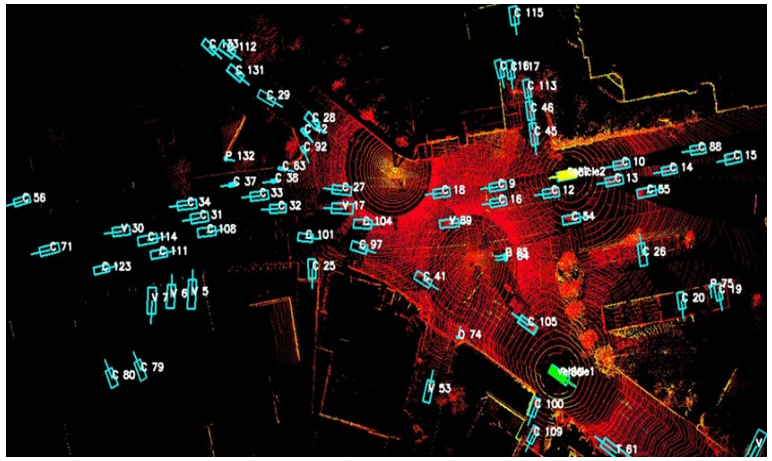
Intersection 1



Intersection 2



Intersection 3



Check out our GitHub repository

GitHub

Check out our website

Website

UrbanIng-V2X: A Large-Scale Multi-Vehicle, Multi-Infrastructure Dataset Across Multiple Intersections for Cooperative Perception

Supported by

HIGHTECH Agenda Bayern

Bundesministerium für Verkehr und digitale Infrastruktur

Bayerisches Staatsministerium für Wissenschaft und Kunst

BayWISS-Kolleg Mobilität & Verkehr

References

[1] Qi Chen, Xu Ma, Sihai Tang, Jingda Guo, Qing Yang, and Song Fu. F-cooper: feature based cooperative perception for autonomous vehicle edge computing system using 3d point clouds. In Proceedings of the 4th ACM/IEEE Symposium on Edge Computing, SEC '19, page 88–100, New York, NY, USA, 2019. Association for Computing Machinery.

[2] Runsheng Xu, Hao Xiang, Xin Xia, Xu Han, Jinlong Li, and Jiaqi Ma. Opv2v: An open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication. In International Conference on Robotics and Automation (ICRA), pages 2583–2589, 2022.

[3] Runsheng Xu, Hao Xiang, Zhengzhong Tu, Xin Xia, Ming-Hsuan Yang, and Jiaqi Ma. V2x-vit: Vehicle-to-everything cooperative perception with vision transformer. In Proceedings of the European Conference on Computer Vision (ECCV), 2022.

[4] Yue Hu, Shaoheng Fang, Zixing Lei, Yiqi Zhong, and Siheng Chen. Where2comm: Communication-efficient collaborative perception via spatial confidence maps. In Thirty-sixth Conference on Neural Information Processing Systems (Neurips), November 2022.

[5] Runsheng Xu, Zhengzhong Tu, Hao Xiang, Wei Shao, Bolei Zhou, and Jiaqi Ma. Cobevt: Cooperative bird's eye view semantic segmentation with sparse transformers. In Conference on Robot Learning (CoRL), 2022.

Technische Hochschule Ingolstadt

AImotion Bavaria

CARISSMA Institute of Automated Driving

TUM