Ego-Trajectory Augmentation on Bird's-Eye View Representation Space for Improving Vehicle Planner

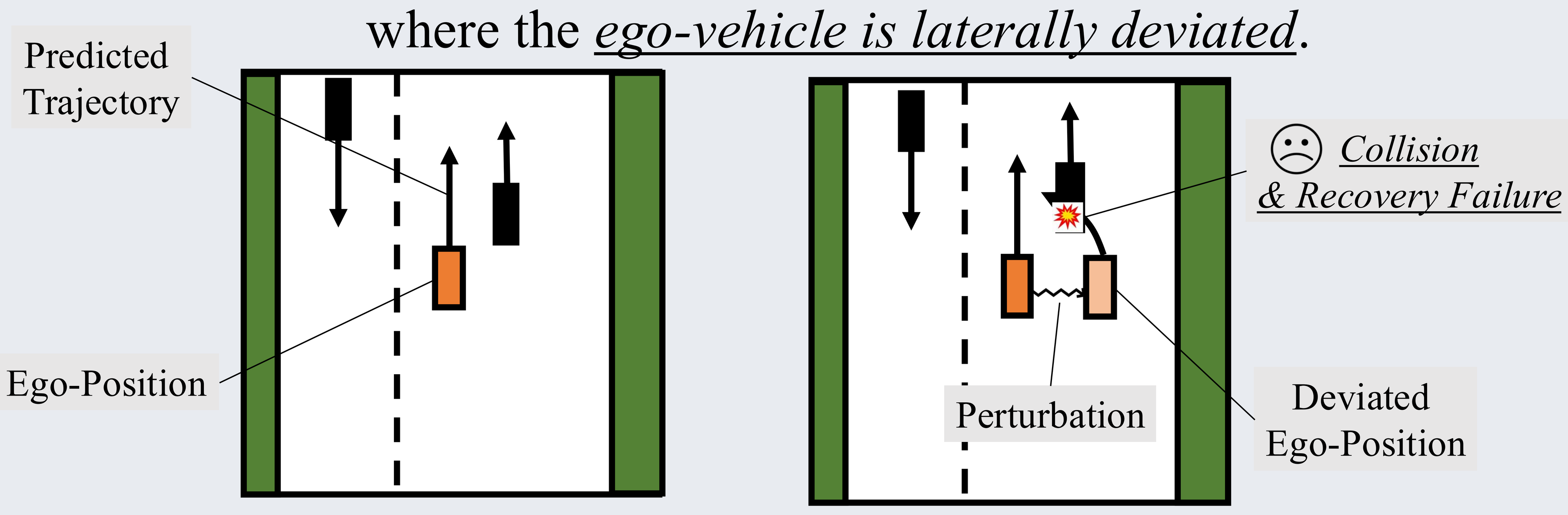
Qiaoyi Deng[1], Yusuke Sekikawa[2], Satoshi Ikehata[2,3], Kenta Hoshino[1,2], Ikuro Sato[1,2]

[1]Institute of Science Tokyo, [2]Denso IT Laboratory, Inc., [3]National Institute of Informatics

1. Introduction

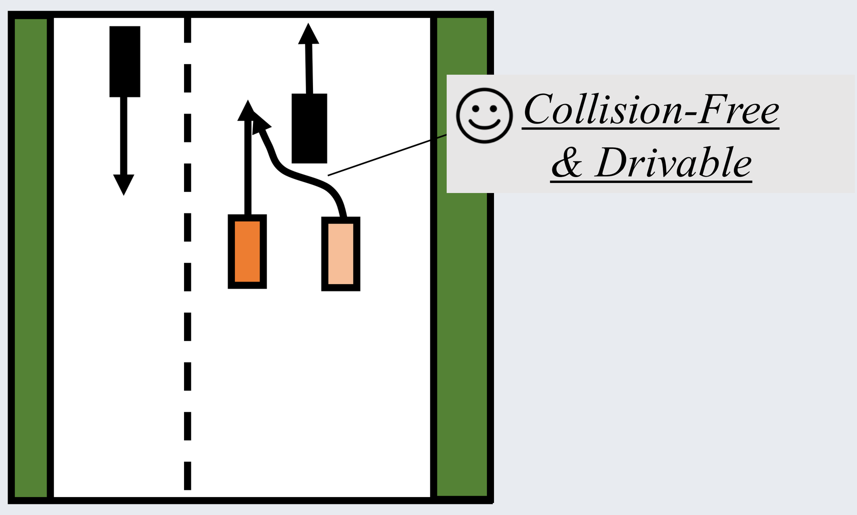
Background End-to-end autonomous driving (E2EAD) models are gaining promising performance in real-world environments.

Challenge Current E2EAD models are vulnerable to situations where the *ego-vehicle is laterally deviated*.



- Potential Risks: *collision* and *recovery failure*

Goal To train the model how to recover from the deviated lateral position without collision in a drivable manner.

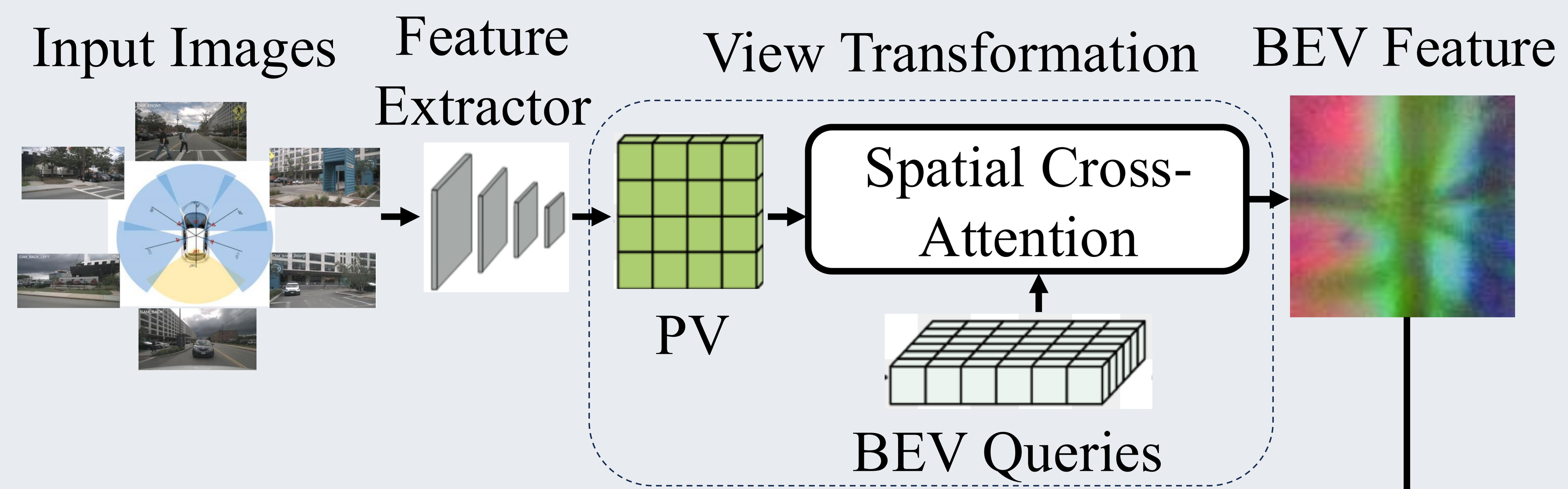


2. Related Work: BEV-Based E2EAD Models

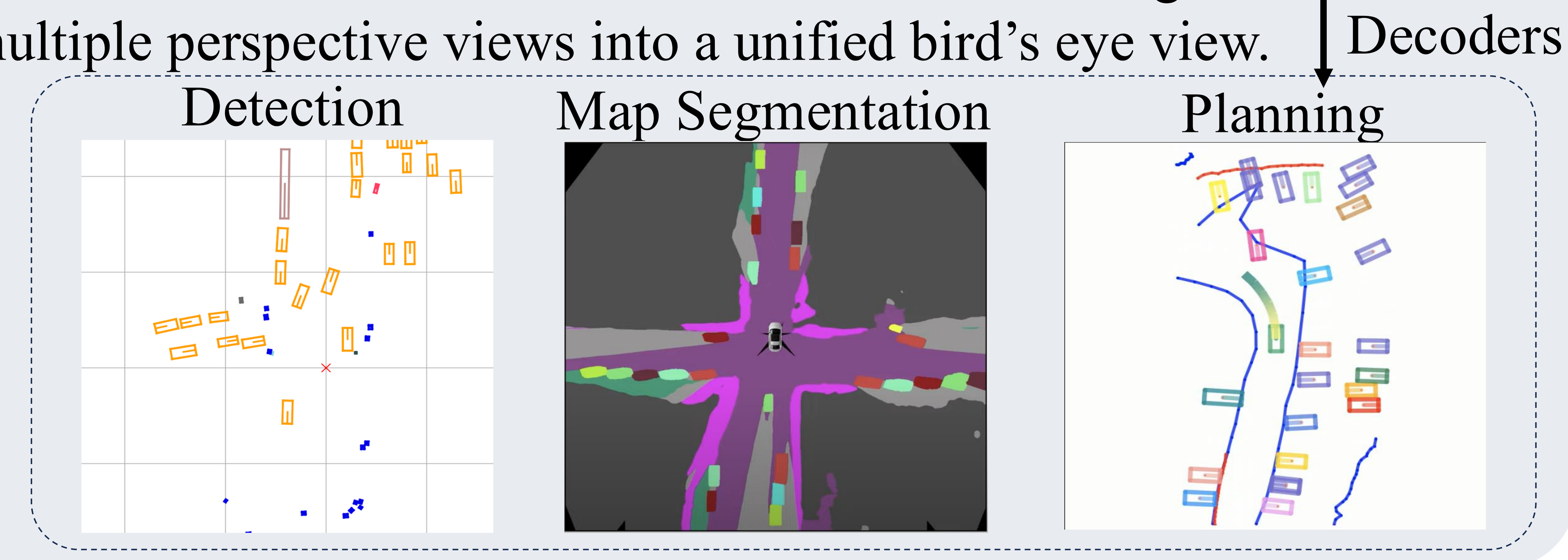
Representative Models

UniAD [Y. Hu, et al., CVPR 2023.] **GenAD** [W. Zheng, et al., ECCV 2024.]
VAD [J. Bo, et al., ICCV 2023.] SSR [P. Li, et al., ICLR 2025.]

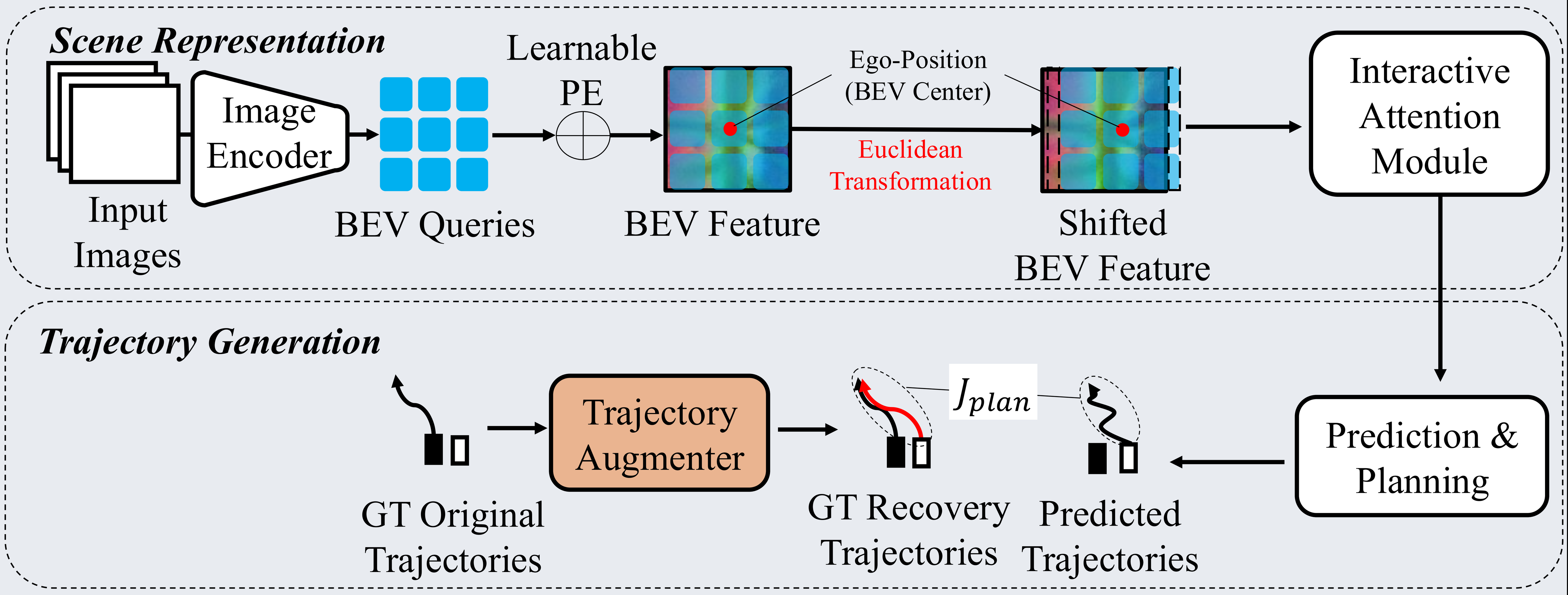
General Framework



Models learn from multi-camera information to integrate multiple perspective views into a unified bird's eye view.



3. Proposed Method



Training Objective: $J = J_{prior} + \lambda_{plan} J_{plan} + \lambda_{map} J_{map} + \lambda_{det} J_{det}$

$$J_{plan} = J_{VAE} + J_{reg} + \lambda_{bound} J_{bound} + \lambda_{col} J_{col} + \lambda_{dir} J_{dir}$$

$$J_{VAE} = D_{KL}(p(z|I), p(z|\Omega)) \qquad J_{reg} = (1-s)\cdot \left\|\tau_{pred} - \tau_{gt}\right\|_1 + s \cdot \left\|\tau_{pred}^{shift} - \tau_{recovery}\right\|_1$$

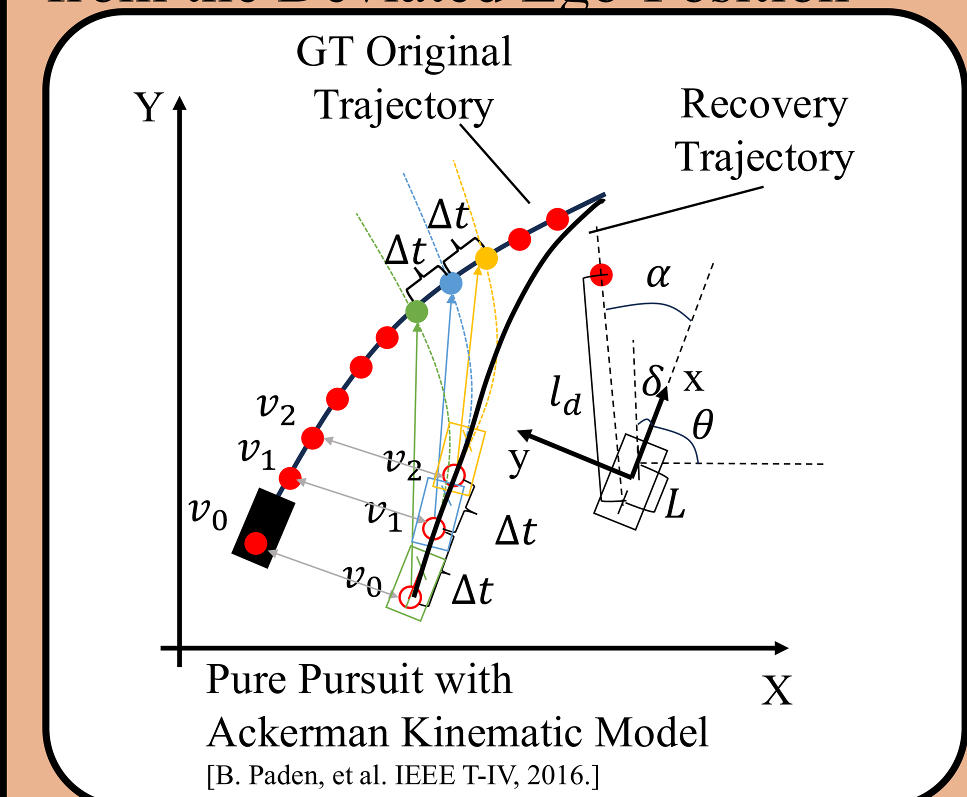
Featue-conditioned Gaussian posterior in latent space — GT Trajectory-conditioned Gaussian prior in latent space

$s \sim Bernoulli(\epsilon)$
ϵ: lateral shift probability

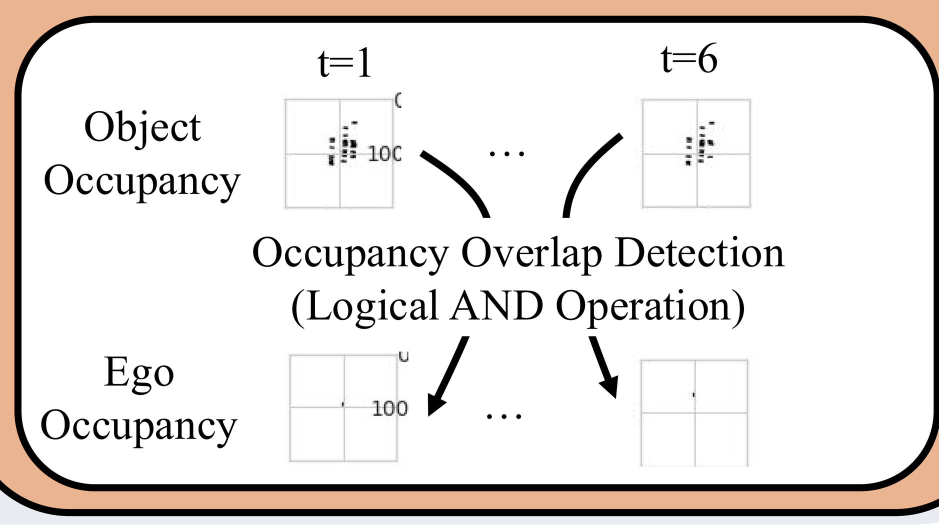
z: Latent vector sampled from distribution representing current trajectory state
I: BEV feature — Ω: Ground Truth ego-trajectory

Trajectory Augmenter

(a) Producing Recovery Trajectories from the Deviated Ego-Position



(b) Collision Detection

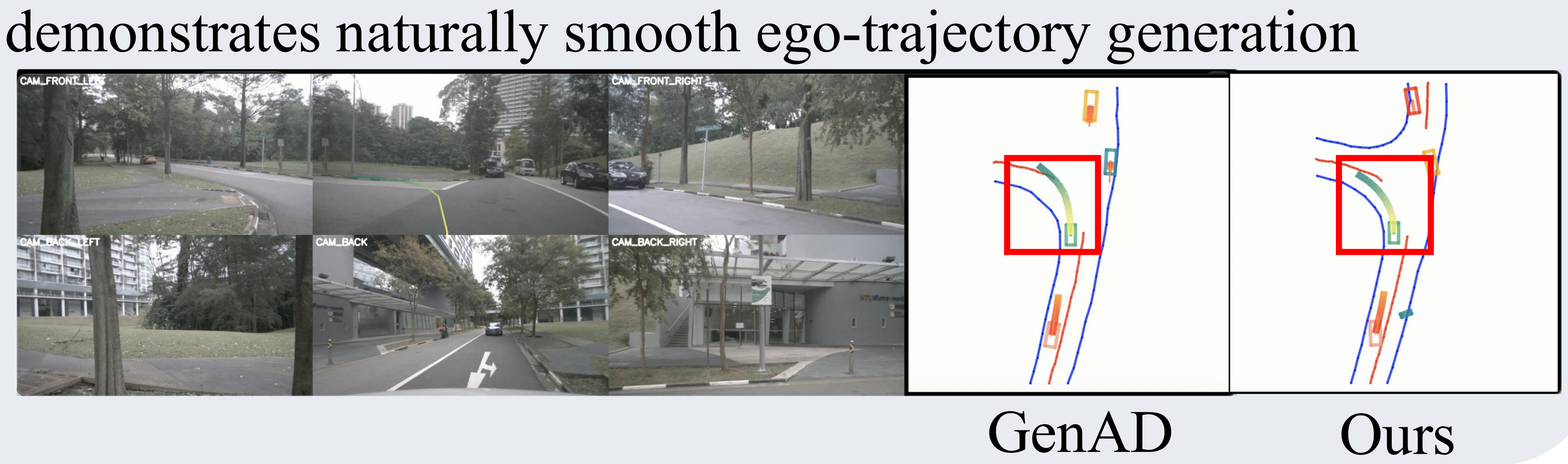


4. Evaluation on the nuScenes dataset [H. Caesar, et al. CVPR, 2020.]

Ours planner outperforms the baseline (GenAD) by 0.4m for L2, 0.2% for CR

Method	L2 (m)↓				Collision Rate (%)↓			
	1s	2s	3s	Avg.	1s	2s	3s	Avg.
GenAD	0.36	0.83	1.55	0.91	0.06	0.23	1.00	0.43
Ours (ϵ=1.0)	0.27	0.48	0.79	0.54	0.40	0.55	0.70	0.55
Ours (ϵ=0.5)	**0.26**	**0.48**	**0.79**	**0.51**	**0.09**	**0.19**	**0.40**	**0.23**

GenAD exhibits abnormal ego-trajectory generation, while Ours demonstrates naturally smooth ego-trajectory generation



This work is supported by DENSO IT LAB Recognition, Control, and Learning Algorithm Collaborative Research Chair (Science Tokyo).