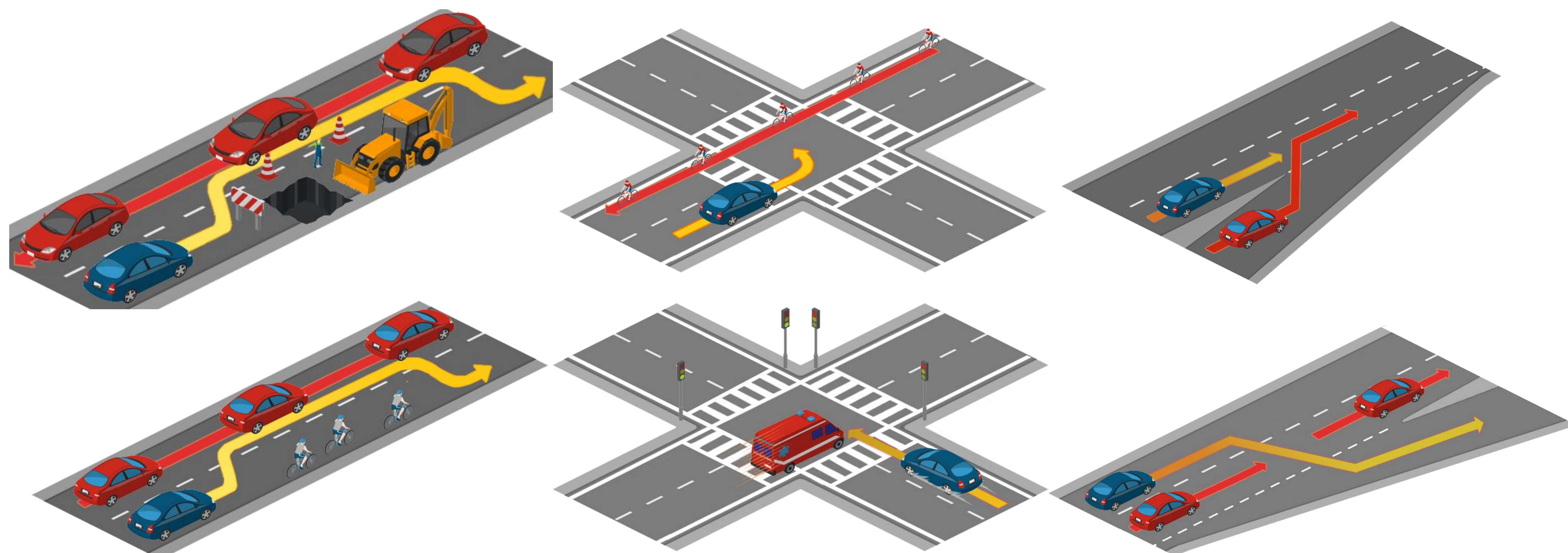


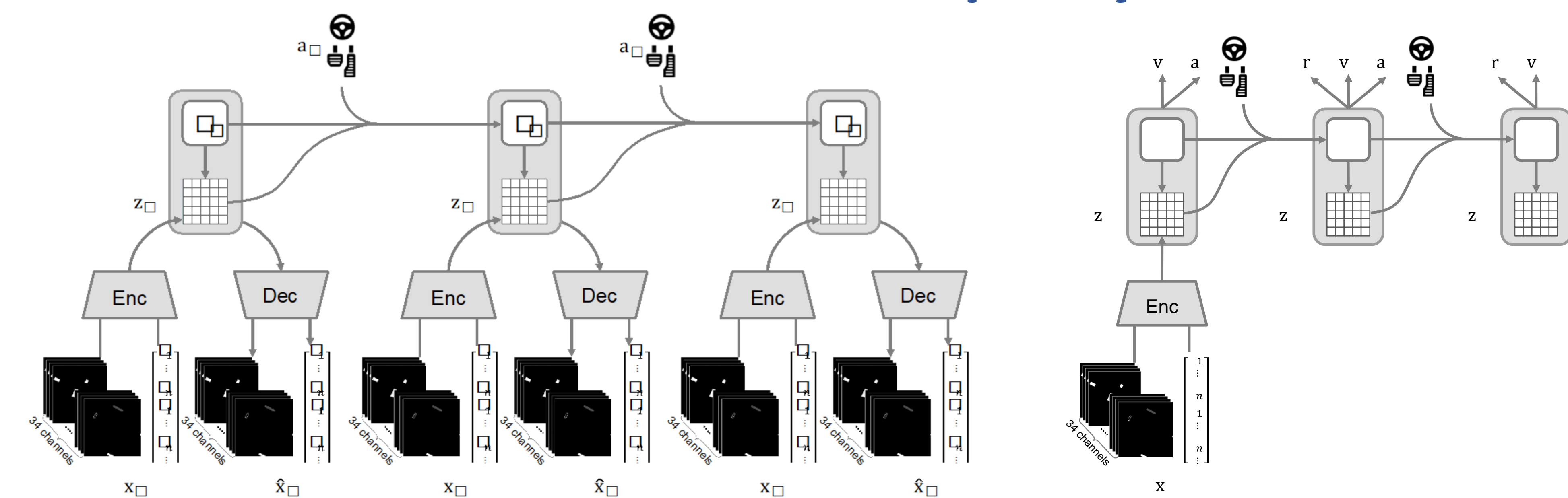
Complex and dynamic traffic scenarios in urban driving



Rule-based planning methods **struggle** to handle these scenarios. Model-based RL, offer hope for urban self-driving, but still encounter 3 main challenges:

1. *Policy degradation* policy is easily to be trapped in the local optima
2. *Long-tail nature* long-tail nature
3. *Vehicle heading stabilization.*

Model the world and learn the policy in the latent space



We model the state transition of the environment with **world model** and make planner **think to drive** in latent space. Gain *thousands of times* acceleration.

Devised Bricks for the urban AD task

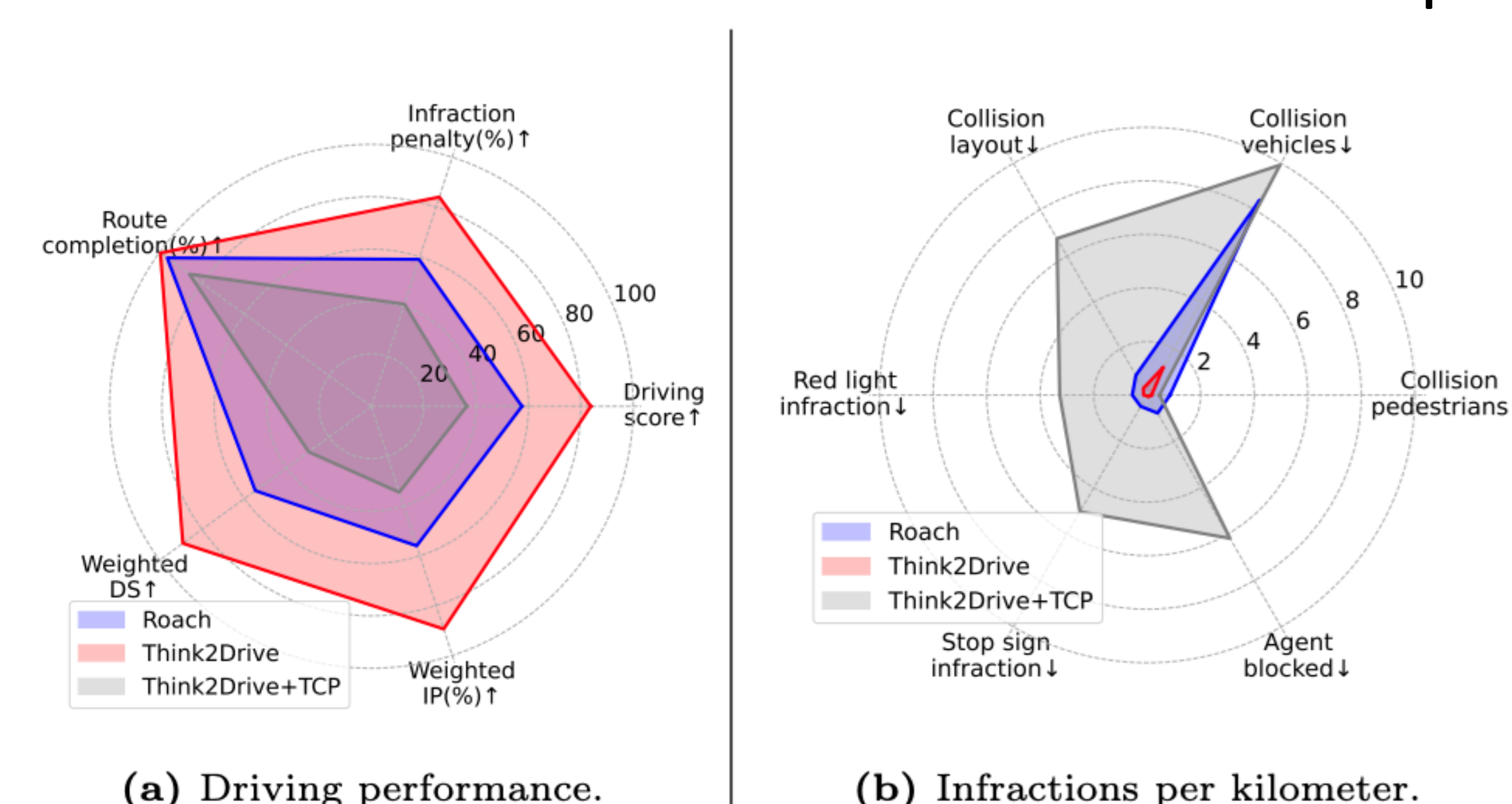
1. Policy degradation -> reset technique
2. Long-tail nature -> scenario generator
3. Sparse valuable transitions -> termination-priority sampling strategy
4. Vehicle heading stabilization -> steering cost function
5. Excessively steep learning curve -> curriculum learning
6. delayed learning signal -> incremental train ratio
7. Excessive exploration -> asynchronous reloading and parallel execution

Experimental results

- Driving performance and infraction of agents on the proposed CornerCaseRepo benchmark.

Input	Method	Driving score	Weighted DS	Route completion	Infraction penalty	Collision pedestrians	Collision vehicles	Collision layout	Red light Infraction	Stop sign infraction	Agent blocked
Privileged Information	Roach [45]	57.5±9	54.8±0.5	96.4±1.1	0.59±0.28	0.85±0.56	8.42±4.65	0.85±0.51	0.56±0.45	0.49±0.44	0.78±0.31
	Think2Drive (Ours)	83.8±1	89.0±0.2	99.6±0.1	0.84 ± 0.01	0.16±0.01	1.2±0.5	0.29±0.02	0.14±0.01	0.03±0.01	0.08±0.01
Raw Sensors	Think2Drive+TCP [42]	36.40±12.23	29.6±0.2	85.88±8.26	0.41±0.32	0.46±0.32	9.92±5.12	6.75±3.08	3.27±1.64	5.03±3.82	6.18±4.62

- Driving performance and infractions on CornerCaseRepo.



- Performance on CALRA Leaderboard V2 official test routes

Method	benchmark	Driving Scores	Weighted Driving Score	Route Complete %
Roach (Expert)	CARLA Leaderboard v1	84.0	-	95.0
Think2Drive (Ours)	CARLA Leaderboard v1	90.2	90.2	99.7
PPO (Expert)	CARLA Leaderboard v2	0.7	0.6	1.0
Think2Drive (Ours)	CARLA Leaderboard v2	56.8	91.7	98.6

- World model prediction

