Taming an End-to-End Autonomous Driving Policy for Urban Navigation of Quadruped Robots

Joochan Kim[1*•] Chanuk Yang[2*] Tackgeun You[2] Ziran Wang[1] Hwasup Lim[2]

[1] Purdue University [2] Korea Institute of Science and Technology (KIST) * Equal contribution • Work done at KIST

IROS 2026 PITTSBURGH

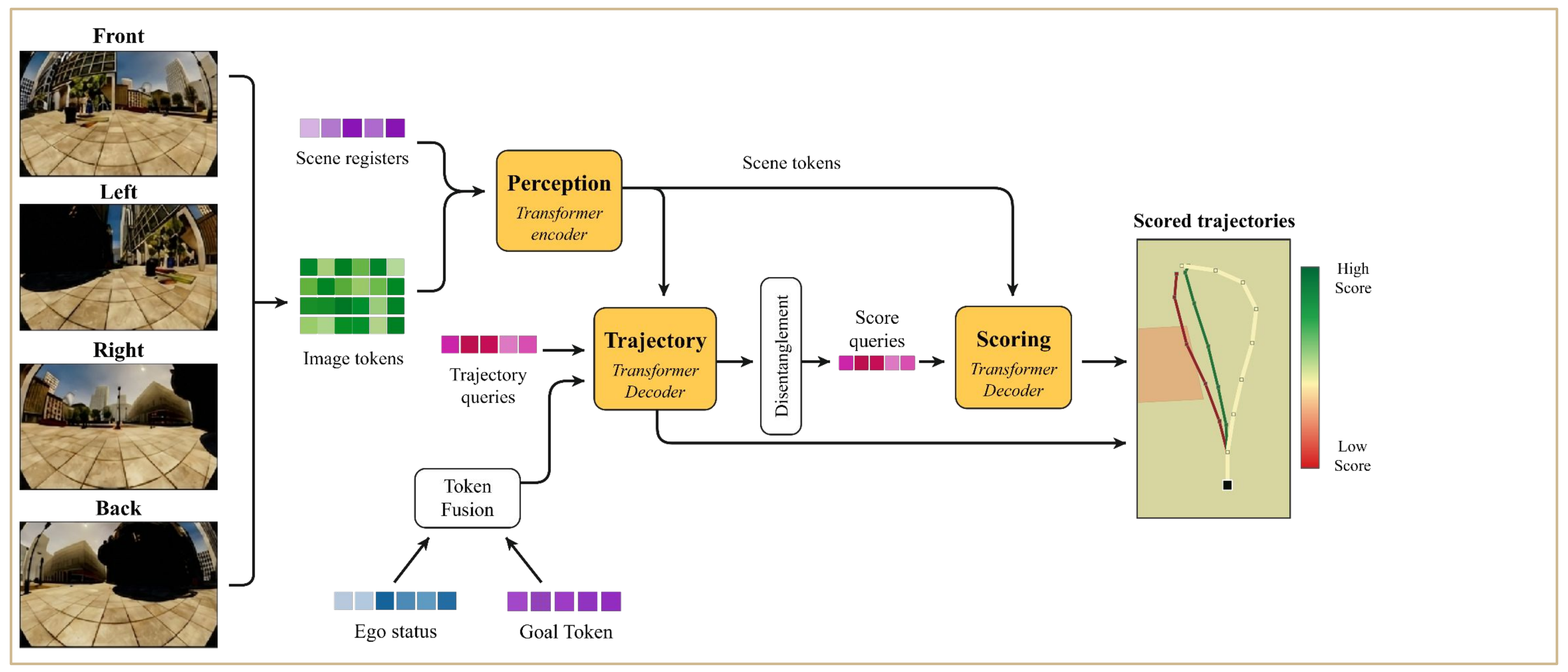
TL;DR Go2-DrivoR adapts the DrivoR end-to-end driving planner to goal-conditioned quadruped navigation with a goal token and sidewalk-aware scoring. Trained only in simulation, it transfers zero-shot to real Unitree Go2 logs.

01 Motivation & Method

Why adapt a driving planner?

- End-to-end driving planners many now perform strongly on benchmarks.
- Quadrupeds cross roads, walkways, and mixed urban spaces with no lane centerlines.
- Vehicle scoring is road-aligned: no preference for sidewalks over roads.
- DrivoR has no explicit goal: no waypoint conditioning, no progress-to-goal term.

Go2-DrivoR architecture



Fisheye views → scene tokens. Fused ego + goal tokens condition 64 trajectory queries; the adapted scorer ranks candidates. DrivoR's decoders are kept unchanged.

What we change

- **Goal token:** 4 s lookahead waypoint, fused with ego state, added to every query.
- **Sidewalk DAC:** rewards sidewalk over road (chips below).
- **Robot footprint** replaces vehicle footprint in NOC / TTC.
- **GEP:** adds progress toward the local goal (box below).
- **DDC removed:** quadrupeds follow no lane direction.

DAC 1.0 all on sidewalk | **DAC 0.5** touches road | **DAC 0** out of bounds

$$\text{GEP} = (1 - \alpha)\cdot\text{EP} + \alpha\cdot\text{GP}, \quad \alpha = 0.5$$

- **EP**: DrivoR's original ego-progress target.
- **GP**: how far the trajectory endpoint moves toward the goal, i.e. displacement along the goal direction ÷ min(‖g‖, 4 m), clipped to [0, 1].
- **Why blend**: EP-only and goal-only both do worse than α = 0.5 (Table III).

Training & inference

- **Sim-only:** 14,192 TartanGround samples; no real-world fine-tuning.
- **Fast:** 10.8 ms per sample at batch size 1 (NVIDIA H200).

02 Quantitative Results

Simulation · TartanGround

Split	Method	NOC ↑	DAC ↑	GEP ↑	Score ↑
Seen	Constant Vel.	0.873	0.954	0.534	0.695
	iPlanner	0.898	0.984	0.769	0.734
	Go2-DrivoR	**0.940**	**0.987**	**0.788**	**0.837**
Unseen	Constant Vel.	0.897	0.903	0.596	0.743
	iPlanner	0.937	0.984	**0.783**	0.772
	Go2-DrivoR	**0.958**	**0.989**	0.776	**0.851**

- Best aggregate score on seen and unseen maps; iPlanner has slightly higher unseen GEP.

Real world · Zero-shot

Scene	Method	ADE ↓	FDE@4s ↓	Head. MAE ↓
Pedestrian	Constant Vel.	0.785	1.470	8.94°
	Go2-DrivoR	**0.200**	**0.209**	**5.79°**
Road	Constant Vel.	0.824	1.542	9.22°
	Go2-DrivoR	**0.213**	**0.186**	**6.81°**

- Better on every metric in both scenes; the largest gain is at FDE@4s. Distances in m.

Ablation · goal token & progress blend

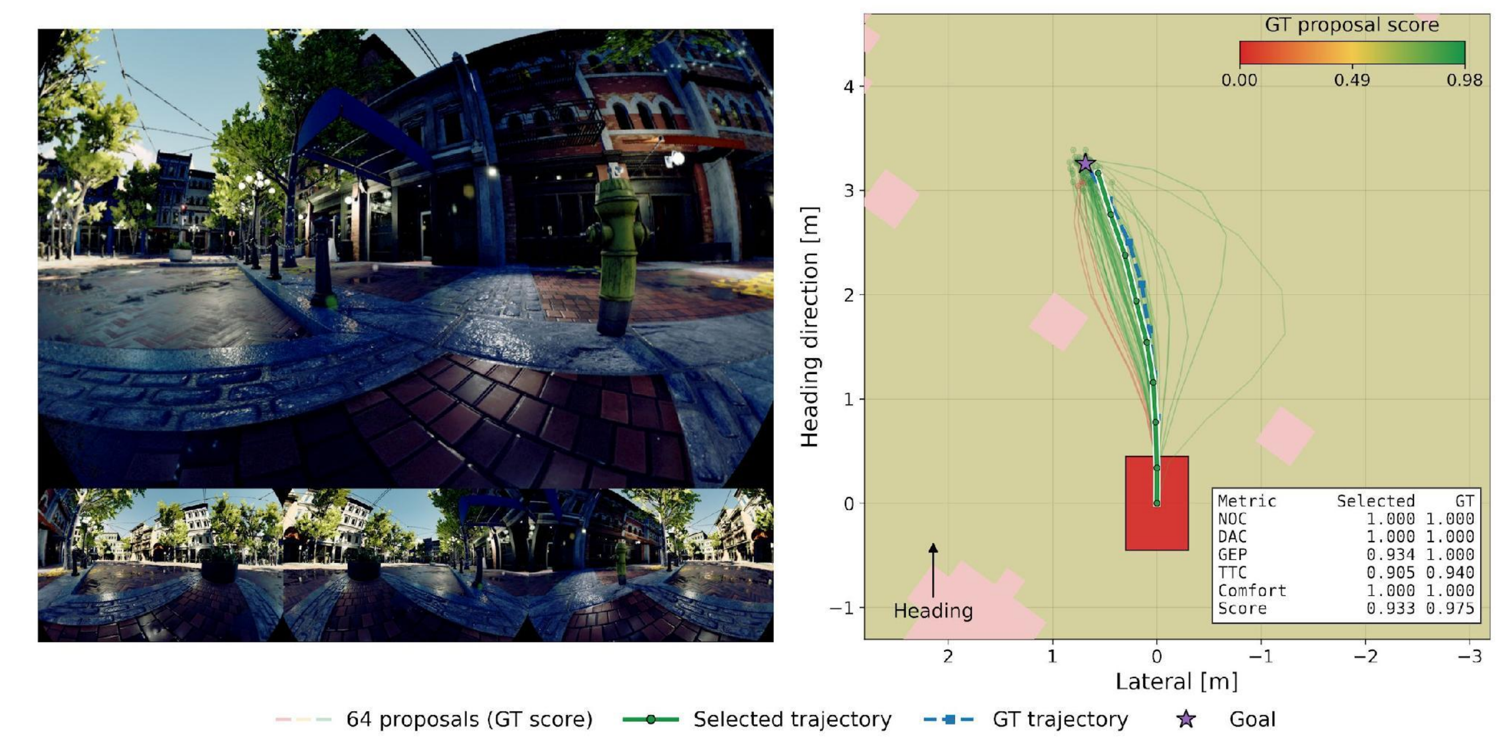
Variant	Goal token	α	ADE ↓	FDE@4s ↓	GP ↑
EP only	✓	0.0	0.251	0.239	0.979
w/o goal token	✗	0.5	4.364	7.766	0.718
Go2-DrivoR	✓	**0.5**	**0.202**	**0.207**	**0.988**
Goal only	✓	1.0	0.325	0.463	0.932

- The goal token is essential; blending ego + goal progress beats either alone. Evaluated on real Go2 logs, both scenes averaged.

Score is an aggregate, not a success rate. iPlanner: official pretrained checkpoint, front depth input. Real-world results are open-loop.

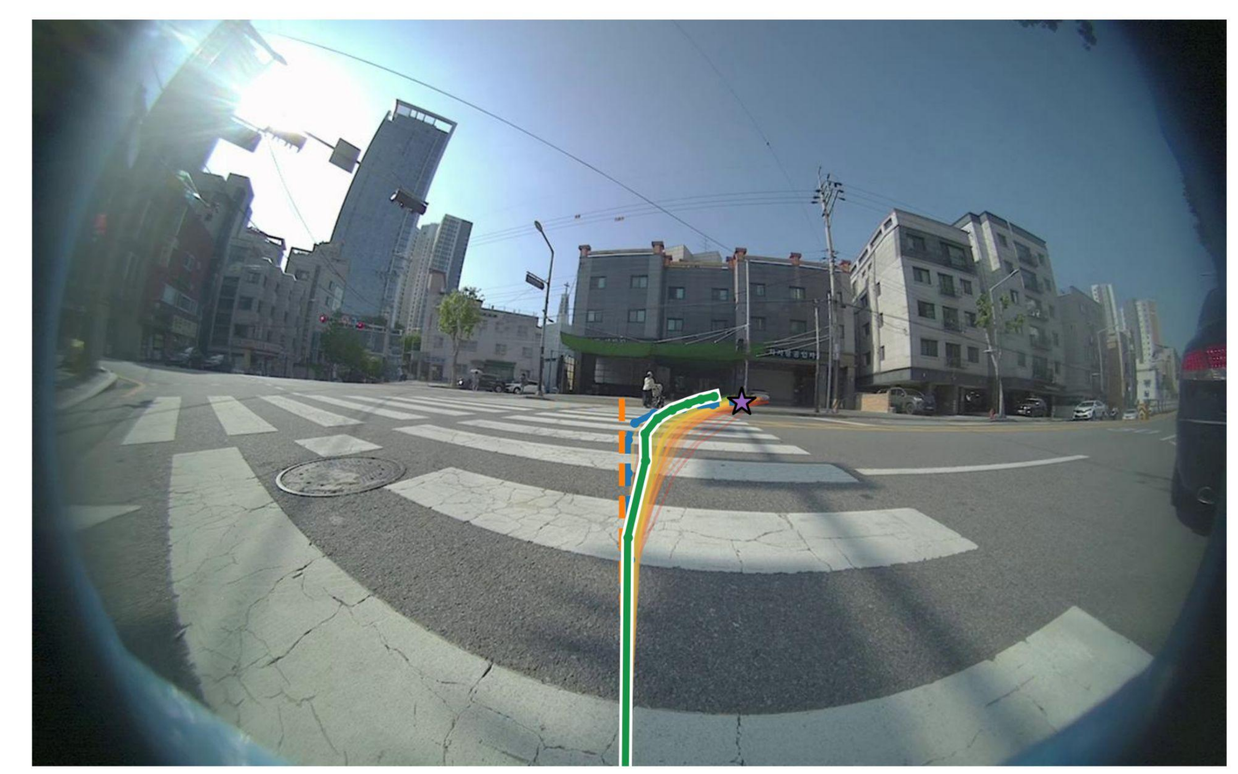
03 Qualitative Results

Simulation



- Selected path reaches the goal and stays in the walkable area.
- Differs from the expert path, yet scores 0.933: a good plan need not copy the expert.

Real world



Green: Go2-DrivoR turns toward the goal (★) **Orange**: constant velocity keeps its heading **Dashed blue**: expert

Takeaways

- Goal token + sidewalk DAC + GEP adapt DrivoR to legged urban navigation.
- Sim-only training → zero-shot, open-loop real-world transfer.

Limitations: expert-derived goals · open-loop only · no dynamic-agent modeling. Next: closed-loop social navigation.

Paper