NAVER LABS Europe | AI FOR ROBOTICS
The 4th NAVER LABS Europe International Workshop on
AI FOR ROBOTICS
20-21 November 2025

RELACS: Reward Learning for Autonomous Driving using Counterfactuals

Eray Çakar[1] Shadi Hamdan[1] Jiazhi Yang[2] Tianyu Li[2] Caojun Wang[2] Hongyang Li[2,3] Fatma Güney[1]

[1] KUIS AI Center, Koç University [2] OpenDriveLab [3] The University of Hong Kong

KOÇ ÜNİVERSİTESİ KUIS

Reward Functions for Driving

Problem: Reliably evaluating driving behavior, particularly in real-world settings, remains a major challenge.

- Traditional reward functions are typically handcrafted, requiring privileged information.
- More recent video-prediction approaches define rewards using likelihood-based metrics and often fail especially in uncertain situations.

Our Approach: Instead, we train a driving reward model directly using observations from CARLA [1]. This approach can potentially scale Reinforcement Learning (RL) to real-world self-driving scenarios by removing the need for manual reward engineering.

Top row: **Expert driving.** *Middle row*: **Crash.** *Bottom row*: **High uncertainty.** *Unlike likelihood-based approaches, our reward model separates uncertainty from actual risk and reliably identifies expert behavior.*

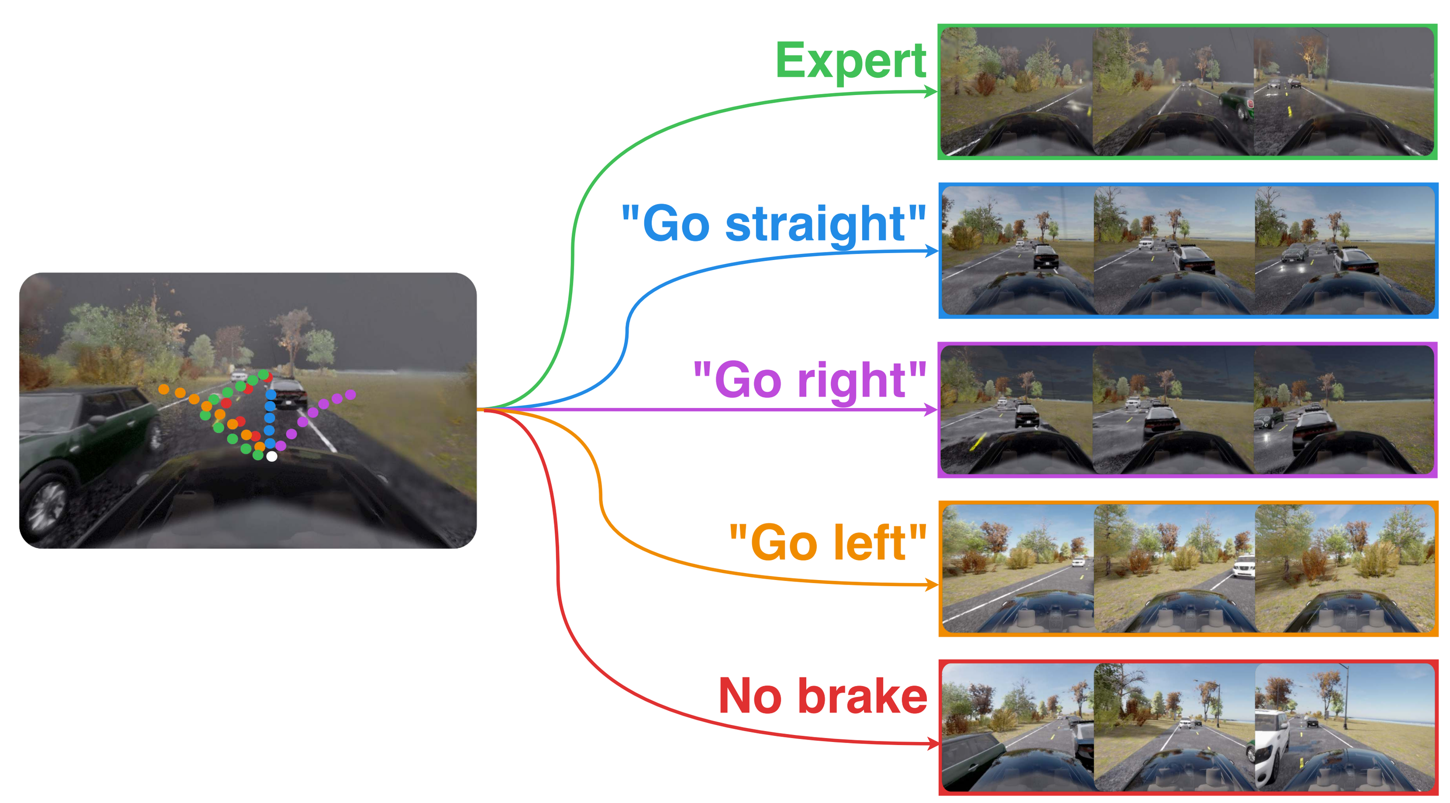
Contributions:

- *COCA:* CARLA driving dataset with counterfactuals.
- *RELACS:* A critical recipe for reward prediction aligned with driving performance.
- *Strong generalization* to both real-world and rendered driving scenarios.

Methodology

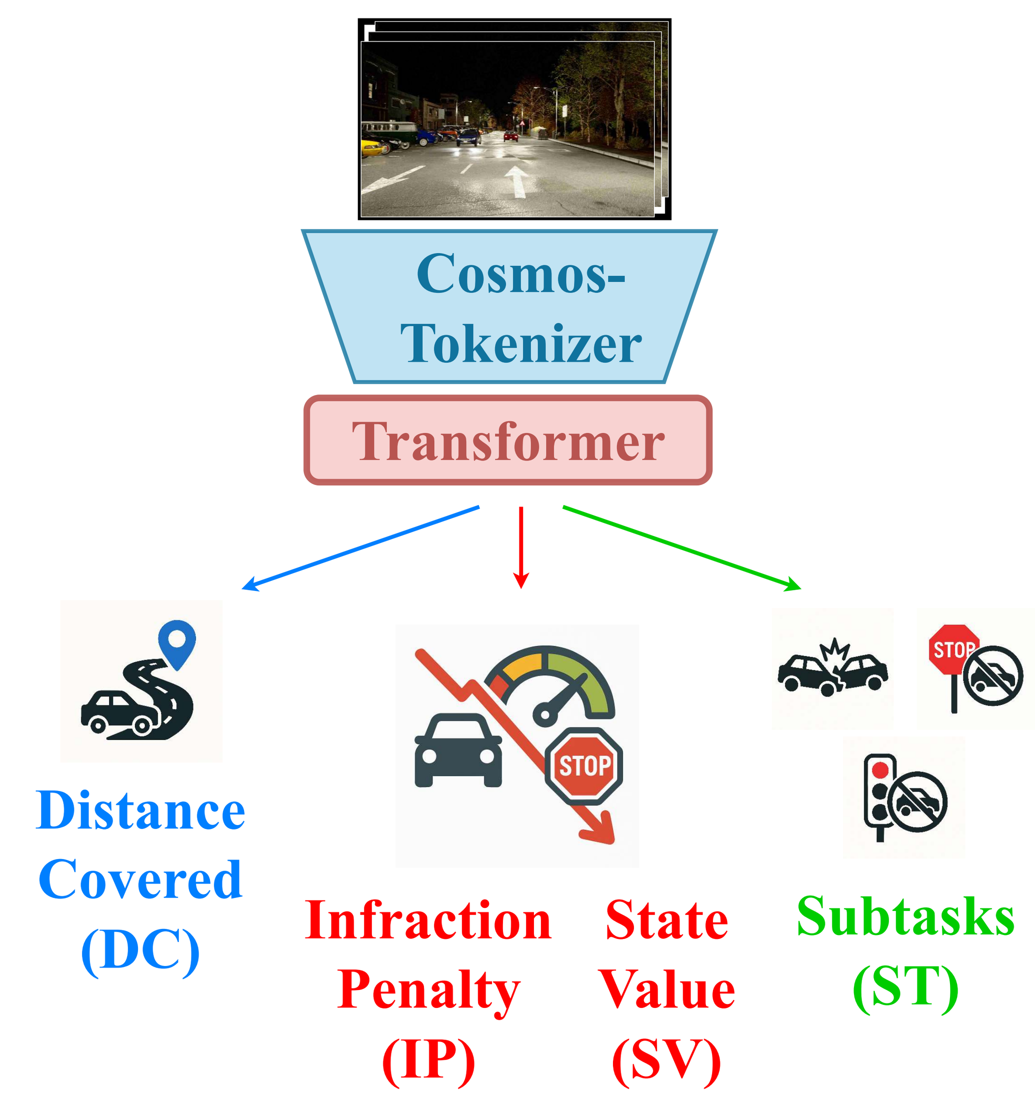
Given a video segment $\mathbf{x} \in \mathbb{R}^{T\times H\times W\times 3}$ of T frames, our goal is to learn a function f that maps the driving behavior observed in $\mathbf{x}$ to a scalar reward r: $f(\mathbf{x}) \mapsto r$.

Counterfactual CARLA (COCA)



To ensure robustness across a wide range of driving behaviors, we construct a dataset of counterfactual driving scenarios using the CARLA simulator.

RELACS



Qualitative Results

Comparison on Youtube driving videos
Top row: **Expert driving.** *Bottom row*: **Suboptimal driving.**

Quantitative Results

Comparison on Youtube driving videos

Label	Metric	Vista[2]	VIPER[3]	RELACS
Expert		0.89	0.35	0.92
Near Crash	Mean	0.88	0.35	0.82
Crash		0.87	0.36	0.54
	$\rho\uparrow$	0.15	-0.25	**0.66**
	$\tau\uparrow$	0.12	-0.20	**0.55**

Comparison on COCA Test set

GT Score	Metric	Vista[2]	VIPER[3]	RELACS
0.0 - 0.2		0.91	0.32	0.37
0.2 - 0.4		0.90	0.36	0.45
0.4 - 0.6		0.90	0.35	0.59
0.6 - 0.8	Mean	0.90	0.33	0.75
0.8 - 1		0.89	0.34	0.89
$\geq$1		0.91	0.39	0.99
	$\rho\uparrow$	0.08	0.11	**0.80**
	$\tau\uparrow$	0.05	0.08	**0.62**

Comparison on nuScenes (left) and NAVSIM (right)

Method	Mean
Vista[2]	0.87
VIPER[3]	0.32
RELACS	0.92

Scenario	Metric	Subset	Value
EP	DC	Slower	0.19
		Real	0.29
		Faster	0.38
RD	SV	Real	0.79
		Rendered	0.61

References

[1] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," 2017.

[2] S. Gao, J. Yang, L. Chen, K. Chitta, Y. Qiu, A. Geiger, J. Zhang, and H. Li, "Vista: A generalizable driving world model with high fidelity and versatile controllability," 2024.

[3] A. Escontrela, A. Adeniji, W. Yan, A. Jain, X. B. Peng, K. Goldberg, Y. Lee, D. Hafner, and P. Abbeel, "Video prediction models as rewards for reinforcement learning," 2023.

Contact Information

Please contact us at: ecakar20@ku.edu.tr
fguney@ku.edu.tr