

CARE Planner : 모방학습 기반 자율주행을 위한 제약 어텐션 및 위험 인지 플래닝

CARE Planner: Constrained Attention and Risk-aware Planning for Imitation-based Autonomous Driving

○김 지 윤¹, 최 경 환^{1*}

¹⁾ 한국과학기술원(KAIST)(E-mail:jiyunkim@kaist.ac.kr, kh.choi@kaist.ac.kr)

Abstract Imitation-learning planners for autonomous driving commonly optimize displacement-based objectives, which improve average trajectory accuracy but may overlook rare unsafe modes. This paper presents CARE Planner, a risk-aware extension of CAR Planner that combines constrained ego-state attention with a Conditional Value at Risk (CVaR)-based tail-risk module. The proposed module estimates clearance-based risk along the prediction horizon, adjusts supervised mode selection toward safer candidates, and constructs risk-aware soft targets for multimodal trajectory learning. The attention constraint prevents excessive dependence on a small subset of ego-state channels, while the CVaR-based module reshapes the output distribution away from high-risk modes. Experiments on the nuPlan test14-random and test14-hard splits show improved open-loop and closed-loop performance, and a pedestrian-waiting scenario analysis shows a reduced high-risk trajectory distribution.

Keywords Autonomous Driving, Imitation Learning, Risk-aware Planning, Attention Constraint, Conditional Value at Risk

1. 서 론

자율주행 모션 플래닝에서 모방학습은 전문가 주행 로그를 활용해 정책을 학습할 수 있어 널리 사용된다. 그러나 평균 변위 오차(Average Displacement Error, ADE)나 최종 변위 오차(Final Displacement Error, FDE)를 줄이는 목적함수만으로는 드문 위험 상황에서의 안전성을 충분히 보장하기 어렵다. 특히 페루프 주행에서는 작은 예측 오차가 누적되어 충돌, 차선 이탈, 불편한 주행으로 이어질 수 있다[7]. nuPlan benchmark는 이를 대규모 로그 기반 페루프 평가로 다루며[1], 최근 연구들은 ego-state shortcut 완화[2], attention 제약[4],[5], 안전 보조 손실[3] 등으로 평균 오차 중심 학습의 한계를 보완하였다. 본 논문은 이러한 한계를 보완하기 위해 CARE Planner를 제안한다. CARE Planner는 CAR Planner의 ego-state attention 제약에 조건부 가치 위험(Conditional Value at Risk, CVaR) 기반 tail-risk 모듈을 결합한다. 제안 모듈은 각 후보 궤적의 시간축 clearance risk를 계산하고, 위험도가 높은 mode의 확률을 낮추는 soft target을 구성한다. 이를 통해 평균 오차가 작은 후보에만 의존하지 않고, 주변 객체와의 안전 여유를 고려한 mode 선택을 유도한다.

2. 제안 방법

2.1 플래너 구조

CARE Planner는 agent history, ego state, local map을 입력으로 사용하는 Transformer 기반 다중 모드

imitation planner이다. Agent history와 lane polygon은 각각 encoder를 통해 token으로 변환된다. Ego state는 (x, y, ψ, v, a, s) 의 6 차원 벡터로 구성되며, single-query attention을 통해 하나의 ego token으로 집약된다[4]. 이후 Transformer encoder가 ego, agent, map token을 처리하고, decoder는 M개의 ego trajectory mode와 logit을 예측한다[2].

2.2 제약 ego-state attention

CARE Planner는 특정 ego-state channel에 대한 attention 집중을 완화하기 위해 Mean-Deviation Constrained Attention(MDCA)을 사용한다[5]. 배치 크기 B, ego-state channel 수 C, n 번째 샘플의 attention vector를 a_{θ}^n 이라할 때, 균등 분포 대비 평균 편차 제약은 다음과 같다.

$$D(a_{\theta}) = \frac{1}{BC} \sum_{n=1}^B \sum_{i=1}^C \left| a_{\theta,i}^{(n)} - \frac{1}{C} \right| \quad (1)$$

$$g(\theta) = D(a_{\theta}) - \delta \leq 0 \quad (2)$$

위 제약은 Augmented Lagrangian penalty로 학습에 반영된다.

$$L_{ALM} = \lambda[g(\theta)]_+ + \frac{\rho}{2}[g(\theta)]_+^2 \quad (3)$$

2.3 CVaR 기반 tail-risk 학습

각 후보 mode k에 대해 ego와 주변 agent 사이의 시간축 clearance risk를 계산하고, CVaR을 이용해 tail risk r_k 를 산출한다[6]. Clearance risk는 ego와 주변 agent 사이의 거리가 작을수록 큰 값을 갖도록 정의한다. z_t^k 를 t 시점의 risk, q_{α}^k 를 α -quantile로 두면 r_k 는 다음과 같다.

$$r_k = \frac{1}{(1-\alpha)T} \sum_{t=1}^T \max(z_t^{(k)} - q_\alpha^{(k)}, 0) \quad (4)$$

Mode probability 를 p_k 라 할 때, risk-aware score 와 best mode 는 다음과 같이 정의된다.

$$s_k = -\log p_k + \lambda_{\text{CVaR}} r_k \quad (5)$$

$$m^* = \underset{k}{\operatorname{argmin}} s_k \quad (6)$$

또한 normalized tail risk 를 이용해 위험도가 높은 mode 의 확률을 낮추는 soft target 을 구성한다.

$$q_k = \frac{p_k \exp(-\lambda_{\text{CVaR}} \tilde{r}_k)}{\sum_{j=1}^M p_j \exp(-\lambda_{\text{CVaR}} \tilde{r}_j)} \quad (7)$$

2.4 학습 목적함수

최종 손실은 imitation loss, risk-aware soft target 을 위한 KL loss, ALM penalty 로 구성된다.

$$L = L_{\text{task}} + \lambda_{\text{KL}} L_{\text{KL}} + L_{\text{ALM}} \quad (8)$$

이를 통해 평균 오차 중심의 mode 선택을 보완하고, 안전 여유를 고려한 trajectory mode 선택을 유도한다.

3. 실험 결과

실험은 nuPlan benchmark 의 test14-random 및 test14-hard split 에서 수행하였다[1]. 평가 지표는 open-loop score(OLS), non-reactive closed-loop score(NR-CLS), reactive closed-loop score(R-CLS)를 사용하였다. 비교 대상은 Base, PlanTF[2], CAR Planner[4], Base+Risk, 그리고 제안한 CARE Planner 이다. Base+Risk 는 attention constraint 없이 CVaR 기반 risk-aware learning 만 적용한 ablation model 이며, CARE Planner 는 attention constraint 와 risk-aware learning 을 모두 적용한 full model 이다.

표 1. nuPlan test14-random 전체 성능 비교

Method	OLS	NR-CLS	R-CLS
Base	86.64	80.01	74.48
PlanTF	86.27	85.23	79.36
CAR	87.67	85.91	78.31
Base+Risk	88.19	86.05	80.36
CARE	89.13	87.48	82.52

표 2. nuPlan test14-hard 전체 성능 비교

Method	OLS	NR-CLS	R-CLS
Base	82.48	65.30	53.11
PlanTF	83.34	70.03	59.83
CAR	86.31	70.81	66.12
Base+Risk	83.92	69.65	59.73
CARE	87.76	72.65	67.17

표 1 과 표 2 에서 CARE Planner 는 두 split 의 모든 평가 모드에서 가장 높은 점수를 보였다. 특히 test14-hard 의 R-CLS 는 CAR Planner 의 66.12 에서 67.17 로 향상되어 어려운 페루프 상황에서 더 안정적인 주행 성능을 보였다. 두 결과를 비교하면, risk-aware learning 만 적용한 경우보다 MDCA 와 함께 결합했을 때 OLS, NR-CLS, R-CLS 가 모두 향상되어 두 모듈이 상호 보완적으로 작용함을 확인할 수 있다.

* 본 연구는 과학기술정보통신부의 재원으로 한국연구재단(NRF)의 지원을 받아 수행되었음(RS-202500554087).

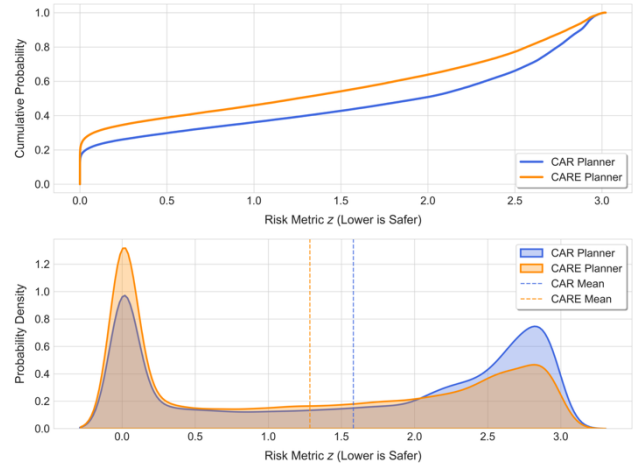


그림 1. test14-hard waiting for pedestrian 시나리오의 risk metric 분포 비교

그림 1 은 test14-hard split 의 waiting for pedestrian 시나리오에서 CAR Planner 와 CARE Planner 의 risk metric z 분포를 비교한 결과이다. 이 시나리오는 ego vehicle 이 보행자의 이동을 기다리거나 저속으로 접근해야 하므로, ego-보행자 간 clearance 변화가 안전성에 직접적인 영향을 미친다. 누적분포함수(CDF)는 특정 risk threshold 이하의 trajectory 비율을, 확률밀도함수(PDF)는 risk 값에 따른 trajectory 분포를 나타낸다. z 값이 낮을수록 안전하며, CARE Planner 는 낮은 위험 구간의 trajectory 비율을 높이고 높은 위험 영역의 밀도를 줄였다. 이는 CVaR 기반 tail-risk 학습이 보행자 대기 상황에서 위험한 후보 궤적의 영향을 완화함을 보여준다.

4. 결론

본 논문은 모방학습 기반 플래너의 평균 오차 중심 학습이 드문 위험 모드를 충분히 억제하지 못하는 문제를 해결하기 위해 CARE Planner 를 제안하였다. CARE Planner 는 ALM 기반 ego-state attention 제약과 CVaR 기반 tail-risk soft target 을 결합하여 입력 표현의 견고성과 출력 분포의 안전성을 함께 개선한다. nuPlan 실험에서 제안 방법은 기존 baseline 대비 open-loop 및 closed-loop 성능을 일관되게 개선하였다. 향후 연구에서는 장면 의미 정보를 활용한 위험 prior 와의 결합으로 확장할 예정이다.

참고문헌

- [1] H. Caesar et al., "nuPlan: A closed-loop ML-based planning benchmark for autonomous vehicles," arXiv:2106.11810, 2021.
- [2] J. Cheng et al., "Rethinking imitation-based planners for autonomous driving," Proc. IEEE ICRA, pp. 14123-14130, 2024.
- [3] J. Cheng, Y. Chen, and Q. Chen, "PLUTO: Pushing the limit of imitation learning-based planning for autonomous driving," arXiv:2404.14327, 2024.
- [4] J. Kim and K. Choi, "CAR Planner: Constrained-attention-based robust imitation learning for autonomous driving," TechRxiv, 2025.
- [5] J. Kim and K. Choi, "Mitigating attention collapse via mean-deviation constrained optimization," TechRxiv, 2025.
- [6] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," Journal of Risk, vol. 2, pp. 21-42, 2000.
- [7] S. Ross, G. Gordon, and D. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," AISTATS, pp. 627-635, 2011.