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

Jiyun Kim * **Kyunghwan Choi** **

* *Gwangju Institute of Science and Technology, Republic of Korea (jiyun6606@gm.gist.ac.kr).*
** *Korea Advanced Institute of Science and Technology, Republic of Korea (kh.choi@kaist.ac.kr)*

Abstract: Most imitation learning planners for autonomous driving are trained with displacement-based objectives such as average displacement error (ADE), which favor average accuracy but overlook how often predicted trajectories become unsafe. CARE Planner augments an attention-constrained Transformer with a Conditional Value at Risk (CVaR) based risk module that measures clearance-based tail risk along the prediction horizon. This risk is used both to select the supervised trajectory mode and to construct a tail-risk-aware soft target that downweights unsafe modes during multimodal learning. Compared to state-of-the-art planners such as PlanTF and our previous framework CAR Planner, CARE Planner outperforms overall performance and significant improvements in safety-related metrics on the nuPlan benchmark, highlighting the effectiveness of risk-aware training in enhancing the reliability of imitation learning planners.

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

1. INTRODUCTION

Safety-critical motion planning for autonomous driving remains a significant challenge, even with advancements in learning-based methods. Imitation learning (IL) is widely adopted as it trains driving policies by mimicking expert demonstrations [Pomerleau (1988)]. However, matching expert behavior on logged data does not guarantee safety. IL policies are prone to distribution shifts and out-of-distribution (OOD) scenarios, and they suffer from compounding errors [Ross et al. (2011)] and shortcut learning [Wen et al. (2020); Chuang et al. (2022)], where spurious correlations are exploited instead of causal driving cues. Consequently, a planner may imitate expert trajectories on average but still fail in safety-critical situations. This highlights the need for safety-focused motion planning frameworks that explicitly incorporate risk.

Nevertheless, many learning-based planners on large-scale datasets such as the nuPlan benchmark [Caesar et al. (2021)] are still designed and evaluated mainly by average imitation error. Their prediction heads are typically trained with soft targets derived from metrics such as ADE, which emphasize mean accuracy but do not distinguish safe trajectory modes from unsafe ones. Our previous proposed framework, CAR Planner, introduced an Augmented Lagrangian–based constrained attention mechanism over ego-state channels [Kim and Choi (2025a,b)] to reduce brittle dependence on a small subset of ego coordinates and to improve robustness under input perturbations. However, its multimodal head still relies on ADE-based soft targets that ignore tail-risk, so rare but safety-critical events remain under-penalized.

To address this limitation, we propose CARE Planner. CARE Planner integrates constrained ego-state attention with a CVaR-based risk module. The risk module evaluates trajectory modes using time-axis CVaR [Rockafellar et al. (2000)] and adjusts the multimodal distribution with a tail-risk soft target. Meanwhile, the constrained attention mechanism ensures balanced weighting of ego-state information. On the nuPlan benchmark, CARE Planner demonstrates competitive performance and stable behavior in challenging scenarios. This highlights the potential of combining attention regularization with tail-risk modeling for safer motion planning.

- **Risk-aware multimodal soft targets for IL planning:** A time-axis CVaR-based tail-risk measure is used to construct mode-wise soft targets that shift probability mass away from risky modes. The classifier is supervised toward a CVaR-rescored safe mode.
- **Constrained attention prior for risk-aware planning:** Building on the Augmented Lagrangian ego-state attention constraint introduced in our previous work, CAR Planner, CARE Planner uses this as an inductive bias. This ensures attention remains aligned with safety-critical ego features while the risk module reshapes the multimodal distribution.
- **Safety and tail-risk improvements on nuPlan:** On the nuPlan `test14-random` and `test14-hard`

* This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-202500554087).

splits, CARE Planner improves overall scores and safety metrics such as collisions, time-to-collision, drivable-area compliance, and comfort. Scenario-level tail-risk visualizations show lower per-step VaR/CVaR and empirical cumulative distribution function (ECDF) curves shifted toward safer regimes in challenging scenes.

2. RELATED WORK

2.1 nuPlan planners

Rule-based stacks like the Intelligent Driver Model (IDM) [Treiber et al. (2000)] are valued for transparency but struggle with distribution shifts and rare events. This has motivated learning-based planners trained on logged driving data. On nuPlan, UrbanDriver [Scheel et al. (2022)] learns closed-loop policies in vectorized scenes, GC-PGP [Hallgarten et al. (2023)] generates goal-conditioned trajectories on lane graphs, and PDM-Open [Dauner et al. (2023)] combines rule-based stacks with learned scoring in an open-loop setting, while PDM-Closed, PDM-Hybrid, and GameFormer [Huang et al. (2023)] extend these ideas to closed-loop and game-theoretic interaction modeling. Transformer-based planners like planTF [Cheng et al. (2024b)] provide strong nuPlan baselines through an attention-based state dropout encoder (SDE) that mitigates shortcut learning on ego states. However, most nuPlan planners are still trained and selected using displacement-based metrics such as ADE and final displacement error (FDE), which emphasize average accuracy on logged data.

2.2 Attention regularized IL planners

Shortcut learning on ego state features has led several works to explicitly regularize attention in imitation learning planners. PlanTF [Cheng et al. (2024b)] introduces an attention-based state dropout encoder that randomly drops ego channels during training to reduce over-reliance on a few coordinates while retaining the standard imitation objective. PLUTO [Cheng et al. (2024a)] adds auxiliary losses that encourage the planner to exploit safety-critical context and maintain consistency between the planned trajectory and scene geometry instead of collapsing onto spurious inputs. CAR Planner [Kim and Choi (2025a)] formulates ego state attention regularization as a constrained optimization problem solved with an augmented Lagrangian method, improving robustness in hazard scenarios compared to other nuPlan state-of-the-art planners. These methods still treat the multimodal output distribution in a risk-agnostic way, since mode labels and probability shaping are mainly driven by average displacement error, so rare but safety-critical modes in the tail of the loss distribution remain under-penalized.

2.3 Risk aware planning with CVaR

A complementary line of work incorporates explicit risk measures into planning and control. CVaR [Rockafellar et al. (2000)] provides a principled way to quantify tail risk by focusing on the expected loss in the worst portion of the distribution and admits convenient optimization formulations. Risk Averse Imitation Learning (RAIL) [Santara et al. (2017)] integrates a CVaR based objective over trajectory costs into the GAIL framework, and evaluates the learned policies on standard continuous control benchmarks in MuJoCo simulation, where a single agent operates in a relatively low-dimensional state–action space under simplified dynamics. This encourages the policy to avoid rare but catastrophic trajectories while still imitating expert behavior.

In motion planning and model predictive control, related works such as Hakobyan et al. [Hakobyan et al. (2019)] pose planning under uncertainty as a CVaR constrained optimization problem for a quadrotor model navigating in a three-dimensional environment with randomly moving obstacles. Their framework combines an RRT*-based reference planner with a CVaR constrained model predictive controller to enforce bounds on safety risk while accounting for disturbances and model error. These approaches show that CVaR is an effective surrogate for tail risk in continuous control, but they are typically applied to lower dimensional systems or small scale simulation benchmarks and are not directly integrated with large scale multimodal planners or imitation learning architectures.

3. CARE PLANNER

3.1 Planner architecture

Conceptually, CARE Planner builds on our previous framework (CAR Planner), which is based on the planTF Transformer encoder–decoder backbone [Cheng et al. (2024b), Kim and Choi (2025a)], and further augments it with a CVaR-based risk head on top of the ALM-constrained ego state attention module to produce multimodal trajectory outputs. At each decision step, the model ingests agent histories, the current ego state, and a local map snippet, and predicts M future ego trajectory modes together with their logits.

3.2 Inputs and tokenization

The inputs are (i) **Agent histories**, past kinematic states of non-ego agents in a local region, (ii) **Current ego state**, a six-dimensional vector containing the ego position (x, y), heading (yaw), longitudinal speed v, longitudinal acceleration a, and steering angle s, and (iii) **Local map**, nearby lane polygons and their attributes.

Agent histories and map polygons are embedded by dedicated encoders into d dimensional tokens. A positional MLP encodes the last pose or polygon center as $[x, y, \cos\theta, \sin\theta]$ and is additively fused with token features. The ego state is not treated as a sequence. Its six scalar channels are aggregated into a single ego token by a learnable attention module over channels.

3.3 Transformer backbone and decoding

Ego, agent, and map tokens are concatenated and processed by a shared Transformer encoder with L layers and h attention heads, using key padding masks for invalid tokens. This joint encoder models interactions between the ego vehicle, surrounding agents, and road geometry via self attention. From the encoded ego token, a trajectory

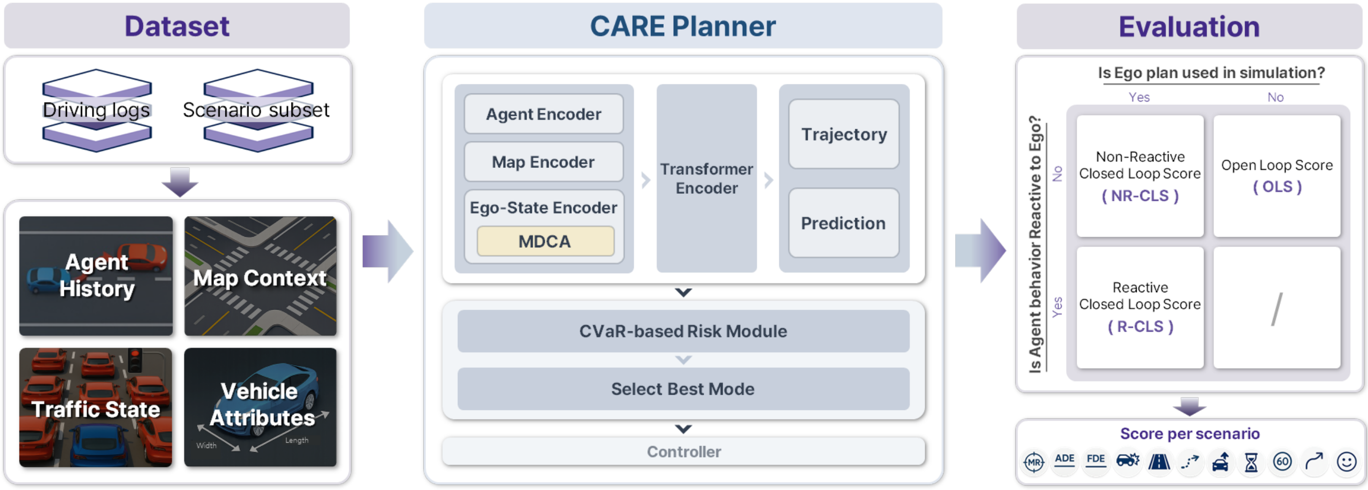


Fig. 1. Overview of CARE Planner. Driving logs are converted into agent, map, and ego features, processed by a Transformer with ALM constrained ego attention and a CVaR based risk head, and the resulting plans are evaluated on nuPlan under OLS, NR-CLS, and R-CLS.

decoder predicts M candidate trajectories of length T and mode logits $\pi \in \mathbb{R}^M$. At test time the most probable mode is selected, headings are reconstructed from $(\cos\psi, \sin\psi)$, and a shallow MLP in parallel predicts future positions for non ego agents.

3.4 ALM constrained ego state attention

The ego state encoder follows the single query attention design of CAR Planner [Kim and Choi (2025b)]. A learnable query attends over C ego channels to produce the ego token, yielding an attention vector $a_\theta \in \Delta^{C-1}$ $(C = 6)$ that specifies how strongly each channel contributes to the representation. CARE Planner reuses the dispersion constraint on this vector: the mean absolute deviation from the uniform distribution must lie below a margin m, and this inequality is enforced with an Augmented Lagrangian Method. This constrained attention mechanism, which we previously termed Mean-Deviation Constrained Attention (MDCA) [Kim and Choi (2025a)].

3.5 CVaR based risk module

As shown in Fig. 2, the CVaR-based risk module is attached on top of the multimodal head. For each sample, the M candidate ego trajectories and the predicted non-ego trajectories are used to compute a per-frame clearance sequence for each mode, based on ego–agent distance relative to an effective radius. This clearance is converted into a nonnegative risk sequence along the horizon, and a time-wise CVaR-style tail risk r_k is obtained for each mode.

These tail risk values are used in two ways. First, the mode index for regression and label construction is chosen by minimizing a risk-augmented score that combines imitation likelihood and tail risk, encouraging the selected mode to remain safe along the horizon. Second, the set of tail risks $\{r_k\}_{k=1}^{M}$ is converted into a tail-risk-aware soft target over the modes, which downweights high-risk modes even when their average error is small. This soft target shapes the mode probability head via a KL loss and biases the

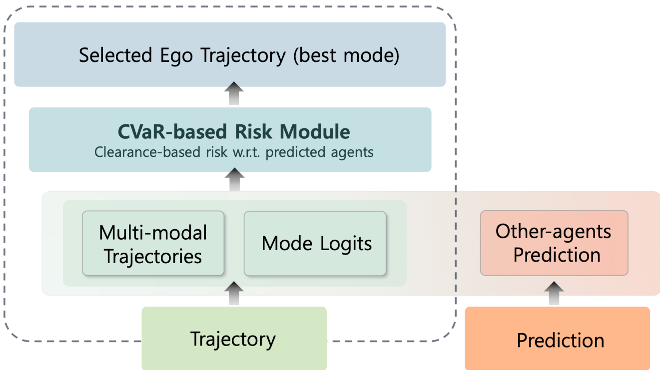


Fig. 2. Overview of the CVaR based risk module attached to the multimodal decoder.

multimodal distribution toward safer trajectory hypotheses while remaining compatible with existing nuPlan-style planners. The computation of these quantities is described in the *Per-frame clearance risk*, *Time-wise tail risk*, and *Risk-aware mode scoring and soft targets* subsections.

3.6 Objectives

CARE Planner follows the pure imitation learning setting of planTF and CAR Planner. For each scene with an expert ego trajectory Y and non-ego trajectories P^{gt}, the planner predicts multimodal ego futures $\{\hat{Y}^{(k)}\}_{k=1}^{M}$, a categorical distribution $p \in \Delta^{M-1}$ over modes, and auxiliary non-ego predictions P. The CAR Planner task loss is reused and augmented with a clearance-based tail risk term that defines risk-aware mode scores and soft targets.

Base imitation objective Let $Y \in \mathbb{R}^{T\times 4}$ be the expert ego trajectory using $(x, y, \cos\psi, \sin\psi)$ and $\hat{Y}^{(k)} \in \mathbb{R}^{T\times 4}$ the k-th predicted mode. Non-ego agents have ground-truth and predicted positions $P^{\mathrm{gt}}, P \in \mathbb{R}^{(A-1)\times T\times 2}$.

Following CAR Planner, a smooth ℓ_1 (Huber) loss

$$L_1^{\text{smooth}}(A,B) = \frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}} \begin{cases} \frac{1}{2}(A_i - B_i)^2, & |A_i - B_i| < \delta, \\ |A_i - B_i| - \frac{\delta}{2}, & \text{otherwise} \end{cases} \tag{1}$$

with threshold $\delta = 1$ is applied element-wise over valid indices $\mathcal{M}$ for both ego and non-ego regression.

Given a mode index m^* (defined below), the imitation loss is

$$L_{\text{task}}(\theta) = L_1^{\text{smooth}}\big(\hat{Y}^{(m^*)}, Y\big) + \text{CE}\big(p, m^*\big) + L_1^{\text{smooth}}(P, P^{\text{gt}}), \tag{2}$$

where $\text{CE}(\cdot,\cdot)$ is the cross-entropy between p and the one-hot label for m^*.

Ego state attention constraint Let $a_\theta^{(n)} \in \Delta^{C-1}$ be the ego state attention over C channels for the n th sample in a batch of size B. The deviation from the uniform distribution is

$$g(\theta) = \frac{1}{BC}\sum_{n=1}^{B}\sum_{i=1}^{C}\left|a_{\theta,i}^{(n)} - \frac{1}{C}\right| - m \le 0, \tag{3}$$

with margin $m > 0$ (for example $m = 0.12$ when $C = 6$). This limits excessive peaking while allowing moderate concentration on informative channels.

The constraint is enforced with an Augmented Lagrangian Method using a nonnegative multiplier λ and penalty parameter $\rho > 0$. With $[z]_+ := \max(0, z)$,

$$L_{\text{alm}}(\theta, \lambda) = \lambda\,[g(\theta)]_+ + \frac{\rho}{2}\big([g(\theta)]_+\big)^2, \tag{4}$$

and after each optimizer step

$$\lambda \leftarrow \lambda + \rho\,[g(\theta)]_+, \tag{5}$$

with λ initialized at zero.

Per-frame clearance risk For each mode $k \in \{1, \dots, M\}$ and step $t \in \{1, \dots, T\}$, let $p_{\text{ego}}^{(k)}(t) \in \mathbb{R}^2$ be the ego position and $p_{\text{obs},i}(t)$ the position of obstacle i, with N_{obs} obstacles. Using an effective radius R_{eff} defined as

$$R_{\text{eff}} = R_{\text{ego}} + R_{\text{obs}}, \tag{6}$$

the clearance to obstacle i is

$$\text{clr}_i^{(k)}(t) = \big\|p_{\text{ego}}^{(k)}(t) - p_{\text{obs},i}(t)\big\|_2 - R_{\text{eff}}. \tag{7}$$

A soft minimum over obstacles gives a smooth approximation of the minimum clearance

$$\text{clr}_t^{(k)} = -\frac{1}{\beta_{\text{clr}}}\log\sum_{i=1}^{N_{\text{obs}}}\exp\big(-\beta_{\text{clr}}\,\text{clr}_i^{(k)}(t)\big), \tag{8}$$

where $\beta_{\text{clr}} > 0$ controls the sharpness. With a safety margin $m_{\text{clr}} > 0$, the per-frame clearance risk is

$$z_t^{(k)} = \text{softplus}\big(m_{\text{clr}} - \text{clr}_t^{(k)}\big) \ge 0, \tag{9}$$

so frames with comfortable clearance contribute nearly zero risk, while near-collision configurations incur large penalties.

Time-wise tail risk For mode k, define the per-frame risk sequence $z^{(k)} = (z_1^{(k)}, \dots, z_T^{(k)})$ and let Z be a random variable whose empirical distribution is given by these T values. Let $q_\alpha(z^{(k)})$ be the empirical α–quantile (Value at Risk, VaR) of $z^{(k)}$ with $\alpha \in (0, 1)$. The CVaR of $z^{(k)}$ is

$$\text{CVaR}_\alpha\big(z^{(k)}\big) = \mathbb{E}\big[Z \mid Z \ge q_\alpha(z^{(k)})\big], \tag{10}$$

that is, the expected risk in the α–tail of the distribution.

In CARE Planner we use the mean excess loss [McNeil et al. (2015)] as a time-wise CVaR-style tail risk:

$$r_k = \frac{1}{(1-\alpha)T}\sum_{t=1}^{T}\max\big(z_t^{(k)} - q_\alpha(z^{(k)}),\, 0\big), \tag{11}$$

which aggregates how strongly and how often the per-frame risks of mode k exceed $q_\alpha(z^{(k)})$ along the horizon. Modes whose risks stay below the VaR level obtain $r_k \approx 0$, so the absence of tail events is directly reflected in the score, while larger values of r_k indicate trajectories that spend more time in high-risk tail regions. For continuous distributions one has

$$\text{CVaR}_\alpha\big(z^{(k)}\big) = q_\alpha\big(z^{(k)}\big) + r_k, \tag{12}$$

so r_k can be interpreted as a mean excess term, focusing the risk module on how severe the tail beyond VaR is rather than on the absolute scale of the quantile.

Risk-aware mode scoring and soft targets Let the decoder output logits $\pi \in \mathbb{R}^M$ and $p = \text{softmax}(\pi) \in \Delta^{M-1}$ with components p_k. An imitation-based score for mode k is

$$s_k^{\text{im}} = -\log p_k, \tag{13}$$

and the risk-aware score is

$$s_k = s_k^{\text{im}} + \lambda_{\text{cvar}}\, r_k, \tag{14}$$

where $\lambda_{\text{cvar}} \ge 0$ trades off imitation likelihood and tail risk. The label mode index is

$$m^* = \arg\min_k s_k. \tag{15}$$

A risk-aware soft target $q \in \Delta^{M-1}$ is obtained by reweighting the predicted probabilities with normalized tail risk. Let $\tilde{r}_k$ be a per batch normalized version of r_k with zero mean and unit variance, clipped for stability. Then

$$q_k = \frac{p_k \exp\big(-\lambda_{\text{cvar}}\,\tilde{r}_k\big)}{\sum_{j=1}^{M} p_j \exp\big(-\lambda_{\text{cvar}}\,\tilde{r}_j\big)}. \tag{16}$$

The probability head is trained with

Method	test14-random			test14-hard		
	OLS	NR-CLS	R-CLS	OLS	NR-CLS	R-CLS
Base	86.64	80.01	74.48	82.48	65.30	53.11
PlanTF	86.27	85.23	79.36	83.34	70.03	59.83
CAR Planner	87.67	84.91	78.31	86.31	69.48	64.64
Base + Risk	88.19	86.05	80.36	83.92	69.65	59.73
CARE Planner	89.13	87.48	82.52	87.76	72.65	67.17

Table 1. Overall score on nuPlan splits.

Method	test14-hard (R-CLS)					
	Collisions	TTC	Drivable	Comfort	Progress	Speed
Base	88.11	81.50	92.64	88.23	72.79	98.02
planTF	85.84	80.88	92.64	93.01	84.55	97.01
CAR Planner	90.63	85.49	94.02	98.16	84.28	98.22
Base + Risk	89.52	83.82	92.27	90.80	81.61	97.17
CARE Planner	92.03	87.55	95.22	99.26	84.55	98.67

Table 2. Sub metrics on nuPlan test14-hard split under R-CLS.

$$L_{\mathrm{KL}} = \mathrm{KL}(q \,\|\, p) = \sum_{k=1}^{M} q_k \log \frac{q_k}{p_k}. \tag{17}$$

Overall objective The overall CARE Planner loss is

$$L(\theta) = L_{\mathrm{task}} + \lambda_{\mathrm{KL}}\, L_{\mathrm{KL}} + L_{\mathrm{alm}}(\theta, \lambda), \tag{18}$$

with λ_{KL} the soft-target weight and L_{alm} the augmented Lagrangian penalty from the attention regularizer.

4. EXPERIMENTS

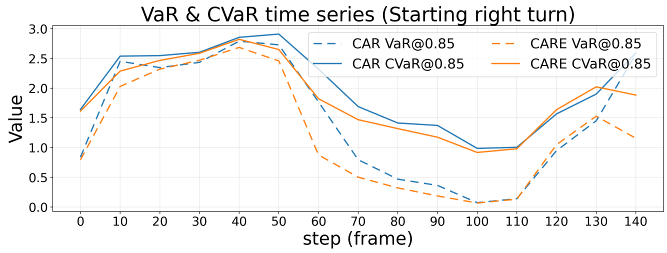
4.1 Experimental setup

All planners are trained for 20 epochs on NVIDIA RTX 4090 GPUs with a batch size of 32, using Adam (learning rate 1e−3, weight decay 1e−4). A state perturbation augmentation adds bounded noise to the ego state and re normalizes coordinates in the perturbed ego frame for compounding errors.

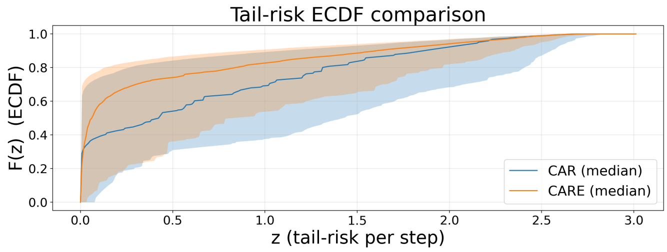
Evaluation follows the nuPlan benchmark on the `test14-random` and `test14-hard` splits, each with 261 scenarios. Reported metrics are OLS (open loop score), NR CLS (non reactive closed loop), and R CLS (reactive closed loop). Higher is better.

The following model variants are compared:

- **Base.** Ego six dimensional state is encoded by a small MLP.
- **Base+Risk.** Base model equipped with the CVaR based risk module that computes clearance based tail risk per mode and uses it for risk aware mode selection and soft targets. No MDCA.
- **PlanTF (with SDE).** Shares the encoder–decoder backbone of Base, but replaces the ego MLP with a single query state attention encoder and State Dropout Encoder (SDE) during training [Cheng et al. (2024b)].
- **CAR Planner.** Shares the encoder–decoder backbone of Base, but replaces the ego MLP with a single query state attention encoder with ego state attention regularized by an augmented Lagrangian dispersion constraint on the six way attention weights [Kim and Choi (2025a)].
- **CARE Planner(Proposed method).** CAR Planner extended with the CVaR based risk module for clearance based tail risk scores and soft targets over modes.



(a)



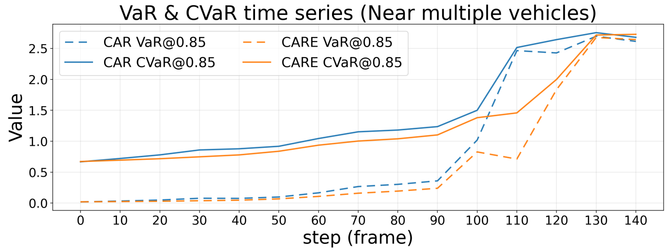
(b)

Fig. 3. Starting right turn visualizations. (a) Timeseries view. (b) Overlay view.

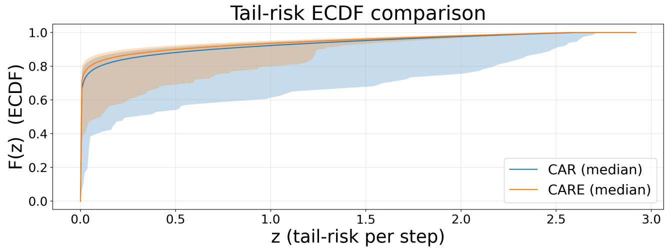
4.2 Experimental results

Table 1 reports the overall nuPlan scores on `test14-random` and `test14-hard`. On `test14-random`, CARE Planner achieves the best performance across all three evaluation modes, improving R-CLS from 74.48 (Base) and 78.31 (CAR Planner) to 82.52 and also slightly raising OLS and NR-CLS. On the more challenging `test14-hard` split, the gap becomes clearer. CARE Planner increases R-CLS from 53.11 (Base), 59.73 (Base+Risk), and 64.64 (CAR Planner) to 67.17. Base+Risk and CARE Planner share the same CVaR-based risk module, so the additional gain mainly reflects the effect of ego-state constrained attention. Compared to CAR Planner, CARE slightly reduces OLS but improves both NR-CLS and R-CLS, indicating a more robust closed-loop behavior on hard scenarios.

Table 2 decomposes the `test14-hard` R-CLS score into safety-related sub-metrics. CARE Planner attains the highest collision score (92.03), time-to-collision (87.55), drivable-area compliance (95.22) and comfort (99.26). Relative to Base+Risk, this shows that adding ego-state constrained attention not only lifts the aggregate R-CLS score but also reduces risky behaviors across individual safety metrics. Compared to CAR Planner, CARE further improves collision and TTC scores while keeping progress and speed-limit compliance at a similar level. These results indicate that CVaR-guided mode selection with soft



(a)



(b)

Fig. 4. Near multiple vehicles visualizations. (a) Timeseries view. (b) Overlay view.

targets, combined with constrained attention, makes the planner more conservative in genuinely risky situations without sacrificing overall driving efficiency.

Figures 3 and 4 illustrate tail-risk behavior on representative scenarios from the `starting_right_turn` and `near_multiple_vehicles` categories. In both cases, the per-step VaR and CVaR time series of CARE stay similar to or below those of CAR, and the ECDF curves in the bottom panels shift consistently to the left. This shift indicates that high tail-risk values occur less frequently under CARE, supporting that the proposed risk-aware supervision produces safer trajectories even within individual challenging scenes.

5. CONCLUSION

CARE Planner augments an attention-constrained Transformer with a CVaR-based tail-risk module for multi-modal imitation planning on nuPlan. Experiments on the `test14-random` and `test14-hard` splits show improved overall scores and consistent gains in safety-related metrics compared to both risk-agnostic and attention-only baselines. These results indicate that combining ego-state attention regularization with time-axis CVaR tail-risk shaping yields more reliable behavior in challenging scenarios.

REFERENCES

Caesar, H., Kabzan, J., Tan, K.S., Fong, W.K., Wolff, E., Lang, A., Fletcher, L., Beijbom, O., and Omari, S. (2021). nuplan: A closed-loop ml-based planning benchmark for autonomous vehicles. *arXiv preprint arXiv:2106.11810*.

Cheng, J., Chen, Y., and Chen, Q. (2024a). Pluto: Pushing the limit of imitation learning-based planning for autonomous driving. *arXiv preprint arXiv:2404.14327*.

Cheng, J., Chen, Y., Mei, X., Yang, B., Li, B., and Liu, M. (2024b). Rethinking imitation-based planners for autonomous driving. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 14123–14130. IEEE.

Chuang, C.C., Yang, D., Wen, C., and Gao, Y. (2022). Resolving copycat problems in visual imitation learning via residual action prediction. In *European Conference on Computer Vision*, 392–409. Springer.

Dauner, D., Hallgarten, M., Geiger, A., and Chitta, K. (2023). Parting with misconceptions about learning-based vehicle motion planning. In *Conference on Robot Learning*, 1268–1281. PMLR.

Hakobyan, A., Kim, G.C., and Yang, I. (2019). Risk-aware motion planning and control using cvar-constrained optimization. *IEEE Robotics and Automation letters*, 4(4), 3924–3931.

Hallgarten, M., Stoll, M., and Zell, A. (2023). From prediction to planning with goal conditioned lane graph traversals. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 951–958. IEEE.

Huang, Z., Liu, H., and Lv, C. (2023). Gameformer: Game-theoretic modeling and learning of transformer-based interactive prediction and planning for autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3903–3913.

Kim, J. and Choi, K. (2025a). Car planner: Constrained-attention-based robust imitation learning for autonomous driving. Institute of Electrical and Electronics Engineers (IEEE). doi: 10.36227/techrxiv.175979328.80125117/v1.

Kim, J. and Choi, K. (2025b). Mitigating attention collapse via mean-deviation constrained optimization. Institute of Electrical and Electronics Engineers (IEEE). doi:10.36227/techrxiv.175877819.93334616/v3.

McNeil, A.J., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press.

Pomerleau, D.A. (1988). Alvinn: An autonomous land vehicle in a neural network. *Advances in neural information processing systems*, 1.

Rockafellar, R.T., Uryasev, S., et al. (2000). Optimization of conditional value-at-risk. *Journal of risk*, 2, 21–42.

Ross, S., Gordon, G., and Bagnell, D. (2011). A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, 627–635. JMLR Workshop and Conference Proceedings.

Santara, A., Naik, A., Ravindran, B., Das, D., Mudigere, D., Avancha, S., and Kaul, B. (2017). Rail: Risk-averse imitation learning. *arXiv preprint arXiv:1707.06658*.

Scheel, O., Bergamini, L., Wolczyk, M., Osiński, B., and Ondruska, P. (2022). Urban driver: Learning to drive from real-world demonstrations using policy gradients. In *Conference on Robot Learning*, 718–728. PMLR.

Treiber, M., Hennecke, A., and Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical review E*, 62(2), 1805.

Wen, C., Lin, J., Darrell, T., Jayaraman, D., and Gao, Y. (2020). Fighting copycat agents in behavioral cloning from observation histories. *Advances in Neural Information Processing Systems*, 33, 2564–2575.