Transferring Causal Driving Patterns for Generalizable Traffic Simulation with Diffusion-Based Distillation

Yuhang Chen[1], Jie Sun[1]*, Jialin Fan[1], Jian Sun[1]

[1]College of Transportation & Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, Shanghai, China
{yuhangchen, jie_sun, jialin_fan, sunjian}@tongji.edu.cn

Abstract

Traffic simulation is essential for validating the safety and reliability of autonomous driving systems, yet data-driven simulation methods often struggle with distribution shifts, limiting their generalizability across diverse datasets (domains). To address this, we present Causal Driving Pattern Transfer (CDPT), a novel two-stage knowledge distillation framework built upon diffusion model to enhance cross-domain generalizability. In Phase I, we implement hybrid self-distillation within the source domain by integrating feature-, response-, and contrastive-level distillation, which enables the model to decompose complex driving behaviors into their core causal components, including scene-conditioned driven patterns, multi-agent interaction dynamics and casual saliency. In Phase II, we introduce a continual distillation strategy: few-shot samples from the target domain are used to initiate generation of diverse synthetic scenarios, allowing the student model to continually adapt to novel environments without retraining on large-scale data. Extensive experiments demonstrate that CDPT achieves strong generalization in both open-loop and closed-loop simulations, effectively generating realistic, interaction-aware behaviors that are critical for scalable and reliable autonomous driving testing.

Project Page —
https://nova-chen151.github.io/simCDPT.github.io/

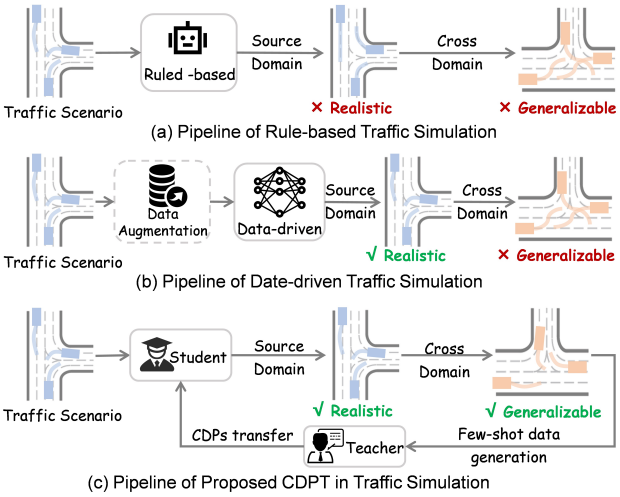

Figure 1: Overview of traffic simulation methods. (a) Rule-based simulation offers interpretable but unrealistic complex interaction behavior with poor target-domain generalization. (b) Data-driven methods can mimic human behaviors but suffer from overfitting and distribution shift. (c) CDPT uses knowledge distillation for causal driving pattern transfer and robust cross-domain generalization.

Introduction

Virtual simulation testing is a cornerstone of autonomous vehicle (AV) validation pipelines, which offers cost-effective, controllable, and scalable evaluation of driving systems under diverse traffic scenarios (Sun et al. 2022; Yan et al. 2023; Wang et al. 2025). By replicating complex, safety-critical conditions that are rare in real-world testing, traffic simulations enable rigorous assessment of AV performance. However, the efficacy of simulation hinges on its ability to generalize across diverse domains, ensuring that synthetic scenarios accurately reflect real-world complexities and remain robust to distributional shifts.

Generalization is critical for simulation fidelity, as AVs must navigate varied environments, from urban intersections to highway merges, often encountering out-of-distribution (OOD) conditions. Traditional rule-based microscopic traffic simulation models, grounded in simplified physical equations and game-theoretic assumptions (Dosovitskiy et al. 2017; Rong et al. 2020), ensure kinematic plausibility but struggle to capture the intricate multi-agent interactions observed in real-world datasets like Waymo Open Motion Dataset (WOMD) (Montali et al. 2023) and Argoverse (Wilson et al. 2023). These limitations have driven the adoption of data-driven approaches, particularly deep learning-based methods such as imitation learning (Fan et al. 2025c; Sun and Kim 2024) and autoregressive models (Zhou et al. 2024; Wu et al. 2024). Although these methods excel at modeling human-like trajectories, they often overfit training distributions, leading to degraded performance in OOD scenarios (Yao, Goehring, and Reichardt 2025).

Diffusion models have emerged as a promising alter-

*Corresponding author.

native, which can generate diverse multimodal trajectories through iterative denoising processes (Guo et al. 2023; Jiang et al. 2024). These models support controlled generation of safety-critical scenarios via guided diffusion (Huang et al. 2024), yet their generalization across datasets remains limited, particularly in multi-agent interactive environments. Techniques like data augmentation and regularization (Monti et al. 2022; Fan et al. 2025a) mitigate these issues partially but fail to address the core challenge: capturing and transferring intrinsic causal driving patterns that govern agent behaviors across domains. These patterns encompass fundamental, context-aware relationships, including scene-conditioned responses (e.g., traffic lights or pedestrian signals triggering braking), multi-agent interaction dynamics, and causal saliency, that govern consistent agent behaviors across varied domains.

To tackle these challenges, we propose *Causal Driving Pattern Transfer (CDPT)*, a novel two-stage self-distillation framework that enhances both simulation fidelity and cross-domain generalization in multi-agent traffic simulation. CDPT leverages the diversity of diffusion models while incorporating a structured knowledge distillation pipeline to extract and transfer Causal Driving Patterns (CDPs)—intrinsic, scene-conditioned causal relationships that explain why agents behave, beyond the merely how they move. Unlike previous work focused on single-agent motion modeling (Monti et al. 2022), CDPT systematically addresses scene-level generalization in interactive, multi-agent environments. Our contributions are as follows.

- We introduce CDPT, a two-stage knowledge distillation framework for traffic simulation using diffusion models, in which we employ feature-, contrastive- and response-based self-distillation to extract transferable CDPs, improving prediction accuracy in open-loop trajectory generation and closed-loop simulation within the source domain (WOMD) in Phase I.
- We propose a continual distillation strategy to enhance cross-domain generalization in Phase II. Using a small batch of target domain data (INTERACTION dataset) to initiate diverse scenario roll-outs via closed-loop simulation, we generate scenarios that preserve causal structures, refining the student model through continual distillation.
- Our distilled model achieves competitive performance in closed-loop simulation in both WOMD and INTERACTION dataset and excels at generating socially interactive, safety-critical scenarios that reflect underlying causal dependencies, demonstrating CDPT's potential for real-world AV testing.

Related Work

Multi-Agent Traffic Simulation

Traditional microscopic traffic models, such as IDM and MOBIL (Treiber and Kesting 2013; Rong et al. 2020), use deterministic kinematic equations and game-theoretic heuristics to simulate vehicle interactions. While ensuring physical plausibility, these rule-based models struggle to capture nonlinear, multi-agent dynamics in complex urban scenarios, producing traffic patterns that deviate from real-world datasets like WOMD (Montali et al. 2023), Argoverse (Wilson et al. 2023), and nuScenes (Caesar et al. 2020).

Recent large-scale trajectory datasets have enabled data-driven approaches for multi-agent simulation. Imitation learning via behavioral cloning replicates expert trajectories but overfits training distributions, leading to cascading errors under distributional shifts (Song et al. 2018; Sun and Kim 2024; Sun and Yang 2024). Transformer-based autoregressive models leverage spatiotemporal attention to model long-range dependencies, improving trajectory forecasting (Zhou et al. 2024; Wu et al. 2024). However, their sequential generation assumes conditional independence, limiting multimodal behavior modeling and causing mode collapse or error accumulation over long horizons (Yao, Goehring, and Reichardt 2025). Diffusion models, which frame trajectory generation as iterative denoising, offer diverse, multimodal synthesis and support guided generation of safety-critical scenarios (Guo et al. 2023; Jiang et al. 2024; Huang et al. 2024). Despite these advances, their reliance on learned distributions hinders cross-domain generalization in interactive multi-agent settings.

Knowledge Distillation for Generalization

Knowledge distillation (KD) transfers latent representations from a teacher to a student model to enhance generalization (Hinton, Vinyals, and Dean 2015). In traffic generative modeling, KD has been applied to transfer driving policies across domains, improve zero-shot motion prediction, and align trajectory representations (Liu et al. 2024; Monti et al. 2022). It has also facilitated map-awareness transfer in scenarios lacking explicit map priors (Wang et al. 2025). However, most KD applications focus on single-agent or non-generative tasks, with limited exploration of scene-level, multi-agent generalization.

To address these gaps, we propose Causal Driving Pattern Transfer (CDPT), a two-stage distillation framework for diffusion-based multi-agent simulation. CDPT integrates feature-, contrastive- and response-based self-distillation to enhance trajectory fidelity and develops continual distillation for cross-domain adaptation, enabling robust, scalable synthesis of interactive traffic scenarios.

Methodology

Problem Formulation

In multi-agent traffic simulation, our objective is to simulate agents' future behavior in dynamic and complex environments. Specifically, we formally define a traffic scenario as $\mathcal{S} = \{\tau, \mathbf{a}, \mathbf{m}\}$, where $\tau = \{\tau^1, \dots, \tau^N\} \in \mathbb{R}^{N \times T \times D_\tau}$ represents the trajectories of N agents over T time steps, with each trajectory state $\tau_n^t \in \mathbb{R}^{D_a}$ including attributes such as position, velocity, and orientation; $\mathbf{a} = \{\mathbf{a}^1, \dots, \mathbf{a}^N\} \in \mathbb{R}^{N \times T \times D_a}$ denotes the driving maneuvers that drive trajectory evolution via the kinematic model $\tau_{t+1} = g(\tau_t, \mathbf{a}_t)$ and $\mathbf{m} \in \mathbb{R}^{D_m}$ encapsulates the environmental context, encompassing the road topology, traffic signal states, and the initial joint state τ_0.

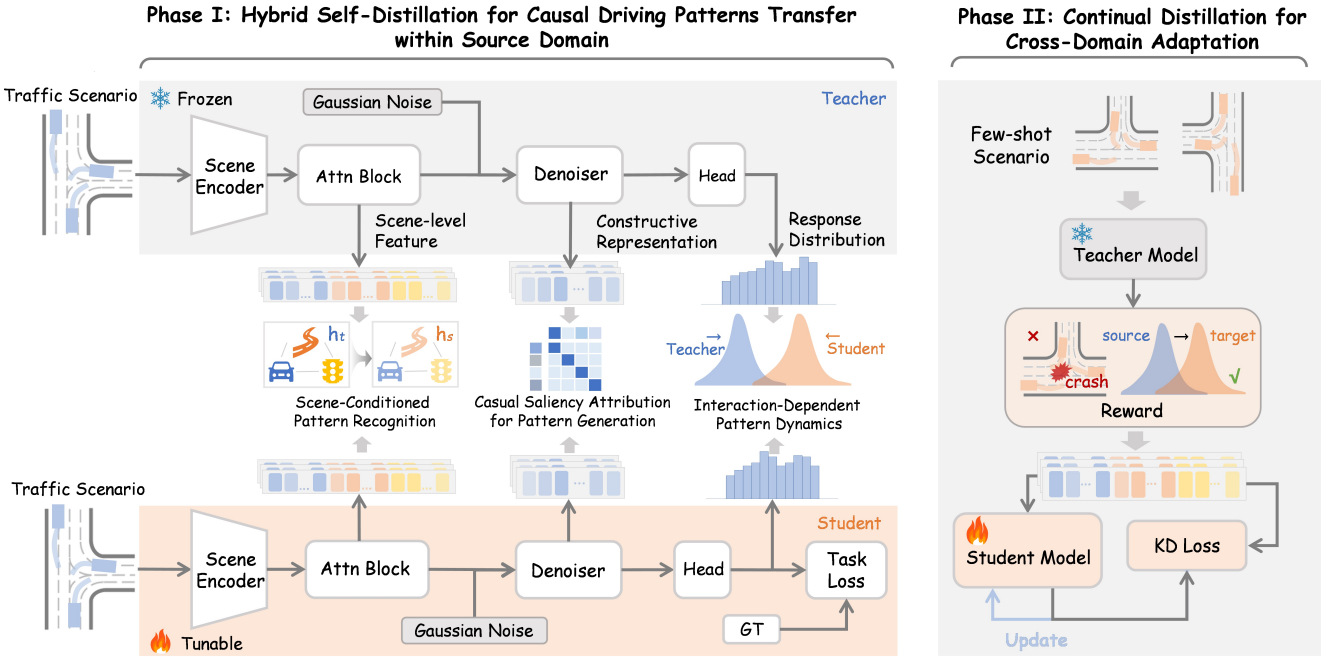

Figure 2: Overview of the CDPT framework for generalizable traffic simulation. **Phase I**: The teacher model encodes causal patterns via diffusion-based generation, and the student learns through feature, response, and contrastive self-distillation. **Phase II**: The teacher synthesizes target-domain scenarios guided by collision-aware and domain-adaptive rewards, enabling continual distillation for student adaptation.

Our primary task is to develop a generative model that samples plausible control sequences $\mathbf{a}$ conditioned on $\mathbf{m}$ to produce trajectories τ consistent with real-world traffic data, reflecting realistic interactive behaviors. Traditional methods typically frame this problem as an optimization task:

$$\min_{\mathbf{v}} \mathcal{L}_\phi(\tau, \mathbf{a}; \mathbf{m}), \tau_{t+1} = g(\tau_t, \mathbf{a}_t), \forall t \in \{0, \dots, T-1\}, \tag{1}$$

where $L_\phi(\tau, \mathbf{a}; \mathbf{m})$ is a parameterized cost function that measures the realism of the scenario. However, designing an effective L_ϕ for complex multi-agent systems remains challenging. Thus, we adopt a data-driven generative approach to directly learn the conditional distribution $p(\mathbf{a} \mid \mathbf{m})$ from real traffic scenario datasets $\mathbb{D}$, enabling natural control sequence generation and subsequent realistic trajectory derivation.

Diffusion Generation for Traffic Simulation

Diffusion models generate realistic multi-agent trajectories by modeling the conditional distribution $p(\mathbf{a} \mid \mathbf{m})$ through a forward noise-adding process and a learned reverse denoising process (Guo et al. 2023; Huang et al. 2024). Specifically, the Denoising Diffusion Probabilistic Model (DDPM) framework (Nichol and Dhariwal 2021) iteratively adds noise to control sequences $\mathbf{a}$ in the forward process:

$$q(\tilde{\mathbf{a}}_k \mid \tilde{\mathbf{a}}_{k-1}) = \mathcal{N}(\tilde{\mathbf{a}}_k; \sqrt{1-\beta_k}\tilde{\mathbf{a}}_{k-1}, \beta_k \mathbf{I}), \tag{2}$$

where $\beta_k \in (0, 1)$ is the noise scale at step k ($k = 1, \dots, K$). Using reparameterization, noisy data is sampled as:

$$\tilde{\mathbf{a}}_k = \sqrt{\bar{\alpha}_k}\mathbf{a} + \sqrt{1-\bar{\alpha}_k}\epsilon, \quad \epsilon \sim \mathcal{N}(0, \mathbf{I}), \tag{3}$$

with $\bar{\alpha}_k = \prod_{i=1}^{k}(1-\beta_i)$. At $k = K$, $\tilde{\mathbf{a}}_K \approx \mathcal{N}(0, \mathbf{I})$, losing structural information.

The reverse process learns to denoise from $\tilde{\mathbf{a}}_K$, approximating the true distribution:

$$p_\theta(\tilde{\mathbf{a}}_{k-1} \mid \tilde{\mathbf{a}}_k, \mathbf{m}) = \mathcal{N}(\tilde{\mathbf{a}}_{k-1}; \mu_\theta(\tilde{\mathbf{a}}_k, k, \mathbf{m}), \sigma_k^2 \mathbf{I}), \tag{4}$$

where the denoiser $\mathcal{D}_\theta$ predicts clean data $\hat{\mathbf{a}}_k = \mathcal{D}_\theta(\tilde{\mathbf{a}}_k, k, \mathbf{m})$, and the mean is:

$$\mu_\theta(\tilde{\mathbf{a}}_k, k, \mathbf{m}) = \frac{\sqrt{\alpha_k}(1-\bar{\alpha}_{k-1})}{1-\bar{\alpha}_k}\tilde{\mathbf{a}}_k + \frac{\sqrt{\bar{\alpha}_{k-1}}\beta_k}{1-\bar{\alpha}_k}\hat{\mathbf{a}}_k, \tag{5}$$

with $\alpha_k = 1 - \beta_k$. The denoiser is trained using the Smooth L1 loss:

$$\mathcal{L}_{\mathcal{D}_\theta} = \min_\theta \mathbb{E}_{\mathbf{a}, \epsilon, k}\left[\mathcal{SL}_1(\mathcal{D}_\theta(\tilde{\mathbf{a}}_k, k, \mathbf{m}) - \mathbf{a})\right]. \tag{6}$$

To stabilize training and capture multimodal trajectories, a behavior predictor estimates trajectory distributions, optimized via:

$$\mathcal{L}_\mathcal{P} = \mathbb{E}_{\mathcal{S}\sim p}[\sum_{n=1}^{N} \mathcal{SL}_1(\hat{\tau}(\hat{\mathbf{a}}^{n,m^*}) - \tau^n) + \gamma CE(m^*, \hat{\omega}^n)], \tag{7}$$

where $\hat{\tau}^n = f(\hat{\mathbf{a}}^{n,m^*})$ are predicted trajectories, m^* is the most probable modality, $\hat{\omega}^n$ is the predicted modality probability, and γ weights the cross-entropy loss. We adopt the VBD (Huang et al. 2024) as our teacher model due to its strong performance in the Waymo Sim Agents Challenge, enabling robust multi-agent trajectory generation.

Two-Stage Distillation for Generalizable Traffic Simulation

Prior traffic simulation methods often prioritize trajectory prediction over understanding causal relationships, limiting generalization across diverse domains. To address this, we propose *Causal Driving Pattern Transfer (CDPT)*, a two-stage distillation framework (Figure 2) that enhances simulation fidelity and cross-domain robustness. **Phase I** employs hybrid self-distillation to capture causal driving patterns in the source domain, while **Phase II** uses few-shot target domain data for continual adaptation.

Phase I: Hybrid Self-Distillation for Causal Driving Patterns Transfer To model complex multi-agent interactions, we introduce a hybrid self-distillation approach combining Feature Self-Distillation (FSD), Response Self-Distillation (RSD), and Contrastive Self-Distillation (CSD). These methods target specific challenges in capturing scene context, interaction dynamics, and causal triggers, enabling realistic and transferable traffic simulation.

Scene-Conditioned Pattern Recognition. Accurate traffic simulation requires understanding scene context (e.g., agent relations, road types, signal states), which prior models often fail to encode systematically, leading to context-agnostic predictions. FSD addresses this by aligning the student's feature representations with the teacher's, ensuring that the model learns to map scene inputs to appropriate driving behaviors. The teacher encoder extracts a scene context vector, and we minimize the discrepancy between normalized feature maps $\mathbf{h}_t$ and $\mathbf{h}_s$:

$$\mathcal{L}_{\mathrm{F}} = \mathbb{E}_{\tilde{\mathbf{a}}_k,\mathbf{m},k}\left[\left\| \frac{\mathbf{h}_t(\tilde{\mathbf{a}}_k,k,\mathbf{m})}{\|\mathbf{h}_t\|_2} - \frac{\mathbf{h}_s(\tilde{\mathbf{a}}_k,k,\mathbf{m})}{\|\mathbf{h}_s\|_2}\right\|_2^2\right], \quad (8)$$

where $\mathbf{h}_t$ and $\mathbf{h}_s$ are the normalized feature representations from the teacher and student encoders at time step k, respectively. This loss ensures the student mimics the teacher's contextual interpretation, improving performance in diverse scenarios.

Interaction-Dependent Pattern Dynamics. Multi-agent interactions involve probabilistic dynamics (e.g., yielding or merging), which prior methods struggle to model due to deterministic or unimodal assumptions. RSD transfers these dynamics by aligning the teacher's and student's action probability distributions using KL divergence:

$$\mathcal{L}_{\mathrm{R}} = \mathbb{E}_{\tilde{\mathbf{a}}_k,\mathbf{m},k}\Big[D_{\mathrm{KL}}\Big(P_T(\hat{\mathbf{a}}^{\mathrm{norm}}_{k,T} \mid \tilde{\mathbf{a}}_k,k,\mathbf{m}) \,\|\, P_S(\hat{\mathbf{a}}^{\mathrm{norm}}_{k,S} \mid \tilde{\mathbf{a}}_k,k,\mathbf{m})\Big)\Big], \quad (9)$$

where $P_T(\cdot)$ and $P_S(\cdot)$ represent the action probability distributions from the teacher and student models, respectively. By minimizing the KL divergence, the student learns not only the most probable action (the peak of the distribution) but also the uncertainty inherent in the decision-making process (the variance of the distribution). This enables the student to capture both the most likely actions and their uncertainties, enhancing realism in interactive settings.

Casual Saliency Attribution for Pattern Generation. Understanding causal triggers (e.g., a pedestrian's movement causing a vehicle to brake) is critical for generating realistic behaviors, but prior methods often ignore these relationships. CSD uses contrastive learning to maximize mutual information between student features and salient triggers, aligning outputs with the teacher's using InfoNCE loss:

$$\mathcal{L}_{\mathrm{C}} = -\frac{1}{B}\sum_{i=1}^{B}\log\frac{\exp(s_{ii})}{\sum_{j=1}^{B}\exp(s_{ij})}, \quad (10)$$

$$s_{ij} = \frac{f_s(\mathbf{z}_i)\cdot f_t(\mathbf{z}_j)}{\|f_s(\mathbf{z}_i)\|_2\|f_t(\mathbf{z}_j)\|_2}, \quad (11)$$

where $f_s(\mathbf{z}_i)$ and $f_t(\mathbf{z}_i)$ are the feature vectors of the student and teacher models for sample i, and $\cdot$ denotes the dot product used to compute cosine similarity. s_{ii} refers to the similarity between the student and teacher outputs for the same sample, while s_{ij} represents the similarity between different sample outputs. This ensures diverse, causally grounded behavior generation by maximizing similarity for positive pairs and minimizing it for negative ones.

Training Objective. To train the CDPT model, we combine task-specific losses ($\mathcal{L}_{\mathcal{D}_\theta}$, $\mathcal{L}_{\mathcal{P}}$) for trajectory accuracy with KD losses ($\mathcal{L}_{\mathrm{F}}$, $\mathcal{L}_{\mathrm{R}}$, $\mathcal{L}_{\mathrm{C}}$) to transfer causal driving patterns, addressing the generalization limitations of prior methods. A simulated annealing operator (λ_{SA}) balances these losses, prioritizing teacher knowledge early and task performance later. The loss function is:

$$\begin{aligned}\mathcal{L}_{\mathrm{CDPT}} = {} & (1-\lambda_{\mathrm{SA}})(\mathcal{L}_{\mathcal{D}_\theta} + \lambda_{\mathcal{P}}\mathcal{L}_{\mathcal{P}}) \\ & + \lambda_{\mathrm{SA}}(\lambda_{\mathrm{F}}\mathcal{L}_{\mathrm{F}} + \lambda_{\mathrm{R}}\mathcal{L}_{\mathrm{R}} + \lambda_{\mathrm{C}}\mathcal{L}_{\mathrm{C}}),\end{aligned} \quad (12)$$

where $\mathcal{L}_{\mathrm{F}}$, $\mathcal{L}_{\mathrm{R}}$, and $\mathcal{L}_{\mathrm{C}}$ are feature, response, and contrastive KD losses for scene context, interaction dynamics, and causal triggers, respectively. The parameter $\lambda_{\mathrm{SA}} \in [0,1]$ dynamically adjusts the trade-off, and $\lambda_{\mathcal{P}}$, λ_{F}, λ_{R}, λ_{C} tune individual loss contributions.

Phase II: Continual Distillation for Cross-Domain Adaptation To generalize to new domains without extensive retraining, we introduce continual distillation (ConD), which uses few-shot target data and the scenario generation capabilities of diffusion model to adapt the student model incrementally. Specifically, the teacher model generates synthetic scenarios aligned with the target domain, guided by a reward function that ensures realism and safety by preventing unrealistic behaviors (e.g., collisions) and aligning with the target domain distribution. Using only a small number of target domain samples, this process refines the parameters of the student model while preserving the causal driving patterns of Phase I, allowing robust cross-domain adaptation.

Reward Function for Scenarios Generation. Generated scenarios must align with the target domain's distribution while avoiding unrealistic collisions, a common issue in diffusion-based models. The reward function combines two components:

- Collision correction. Due to limitations in behavior generation, the teacher model often generates traffic scenarios with a high collision rate. Originally introduced by VBD (Huang et al. 2024), we penalize agent collisions:

$$\mathcal{J}_{\mathrm{col}} = \sum_{t=1}^{T}\sum_{i,j}^{A} d_{ij}(g(\mathbf{a}_t))\,\mathbb{1}\left(d_{ij}(g(\mathbf{a}_t)) < \epsilon_d\right), \quad (13)$$

where $d_{ij}(g(\mathbf{a}_t))$ denotes the Minkowski distance between the footprints of agents i and j at time step t, and ϵ_d is a predefined proximity threshold below which a potential collision is assumed; $\mathbb{1}\left(d_{ij}(g(\mathbf{a}_t)) < \epsilon_d\right)$ is 1 if $d_{ij}(g(\mathbf{a}_t)) < \epsilon_d$, otherwise 0.

- Cross-Domain Adaptation. To align generated trajectories with the target domain, we propose a cross-domain adaptation objective that minimizes the divergence between their statistical distributions. Let $\mathbf{x}^* \sim \mathcal{D}_{\text{gt}}$ denotes trajectories from the target domain and $g(\mathbf{a}_t)$ represents generated trajectories at time t. The adaptation reward is:

$$\mathcal{J}_{\text{cda}} = \mathbb{E}_{\mathbf{x}^* \sim \mathcal{D}_{\text{gt}}}\left[\exp\left(-\frac{\text{DTW}(g(\mathbf{a}_t), \mathbf{x}^*)}{\sigma}\right)\right], \quad (14)$$

where $\text{DTW}(\cdot,\cdot)$ is the dynamic time warping distance, and σ controls matching softness. This encourages realistic, domain-aligned trajectory generation.

The final reward function for data generation is a weighted sum of both components:

$$\mathcal{J}_{\text{T}} = \theta_1 \mathcal{J}_{\text{col}} + \theta_2 \mathcal{J}_{\text{cda}}, \quad (15)$$

where θ_1 and θ_2 are weights to balance the magnitude of the reward functions, prioritizing cross-domain adaptation to align trajectories with the target domain while ensuring collision-free behaviors. Simply, we set $\theta_1 = 1$ and $\theta_2 = 1000$ in our experiments.

Continual Learning with Synthetic Scenarios. Adapting to new domains requires diverse training data, which is often scarce. By generating synthetic scenarios, the teacher model bridges source and target domains, enabling the student to refine parameters without large-scale retraining. This preserves Phase I's causal patterns while enhancing robustness, outperforming traditional methods in capturing diverse traffic behaviors.

Experiments

Experimental Setup

Datasets In this study, we adopt WOMD (Montali et al. 2023) for model training. Due to computational constraints and to accelerate experimental iterations, we selected a subset of WOMD for model training, comprising 48,699 samples in the training set and 44,097 samples in the validation set, which do not hinder the generalizability of the conclusions. To evaluate model performance, closed-loop simulation experiments are conducted via Waymax (Gulino et al. 2023). Furthermore, to rigorously validate the cross-dataset generalization capability, we employ the INTERACTION dataset (Zhan et al. 2019) for further testing. Renowned for its rich scenario diversity and complex driving contexts, the INTERACTION dataset provides a robust benchmark for assessing model adaptability and performance in heterogeneous environments.

Training We follow the training protocol and hyperparameter configurations reported in the VBD (Huang et al. 2024) to ensure comparability. All experiments are conducted on a workstation equipped with 4 NVIDIA A800 GPUs, providing sufficient computational resources for large-batch diffusion model training. All models are trained for 16 epochs in Phase I and 6 epochs in Phase II, with a batch size of 4 throughout.

Metrics To comprehensively assess the performance of multi-agent traffic simulation models, we employ minADE and minFDE as primary evaluation metrics during training in the WOMD. For closed-loop traffic simulation with AV involved, we leverage the Waymax evaluation framework, which quantifies model performance through five core metrics: off-road, collision, wrong-way, kinematic infeasibility and log divergences. These metrics collectively characterize navigation competence, safety compliance, physical plausibility and predictive accuracy in simulated environments.

Results Analyses in Source Domain

Open-loop Trajectory Generation Performance We evaluate open-loop trajectory generation performance using minADE and minFDE in the WOMD validation set. As shown in Table 1, our CDPT achieves a highly competitive minADE and the best minFDE among all the compared methods. Compared to the diffusion baseline VBD, CDPT reduces minADE by 30.7% and minFDE by 25.1%. Against strong autoregressive baselines, CDPT delivers substantially lower minFDE than both SMART and CAT-K while maintaining competitive minADE, despite training on only 50% of the training data used by these baselines. This superior performance with significantly less data highlights the exceptional data efficiency and generalization capability of our continual diffusion pre-training paradigm, establishing new state-of-the-art minFDE in the WOMD validation set under constrained data regimes.

Model	minADE ↓	minFDE ↓
VBD (Huang et al. 2024)	1.199	3.263
SMART (Wu et al. 2024)	0.780	3.620
CAT-K (Zhang et al. 2025)	**0.683**	3.020
CDPT (Ours)	0.831	**2.444**

Table 1: Comparison of methods in WOMD validation set for open-loop trajectory generation performance. Best values are **bolded**

Closed-loop Simulation performance Closed-loop simulation evaluates the model's ability to generate realistic interaction behaviors for AV testing, where AVs replay observed trajectories from the WOMD, and background vehicles (BVs) are simulated to create realistic traffic interactions. We evaluated performance in 914 unseen WOMD scenarios. Table 2 compares methods, with CDPT achieving the best overall performance, particularly in collision, off-road, and log divergence metrics across Phase I and Phase II. CDPT records the lowest collision rates and log divergences, with competitive results in wrong-way and kinematic infeasibility metrics. FSD outperforms the baseline teacher model, notably in collisions and log divergences, but is surpassed by CDPT. In contrast, RSD show higher off road, indicating limitations in dynamic decision-making.

CDPT's minimal performance degradation from Phase I to Phase II underscores its continual distillation strategy, which mitigates catastrophic forgetting and enables robust adaptation to novel scenarios. Figure 3 visualizes the superior rule-compliant behavior of CDPT in complex WOMD scenarios, complementing the quantitative results.

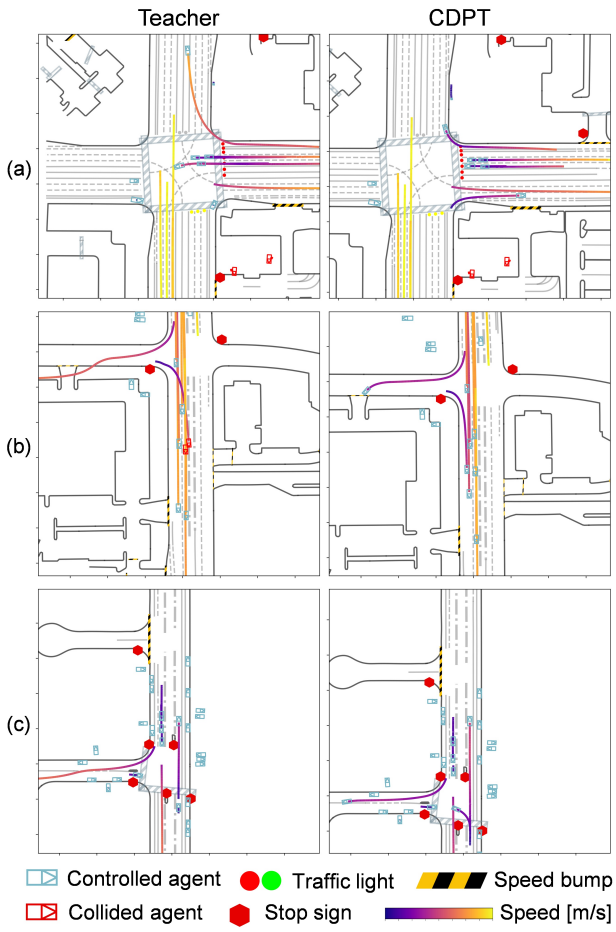

Figure 3: Illustration of closed-loop simulation in WOMD: (a) At a signalized intersection, CDPT respects traffic signals and executes a safe stop and yield maneuver, whereas the teacher model violates the signal phase; (b) In a main-road merge scenario, CDPT performs a smooth and collision-free merging behavior, while the teacher model leads to multi-agent conflicts and infrastructure contact; (c) At an unsignalized T-junction, CDPT generates stable turning and yielding behaviors, correcting the teacher model's static or erratic motions.

Cross-Domain Generalization Results

To evaluate the generalization of the model in diverse traffic environments, we conduct closed-loop simulation experiments in 498 dense, interaction-heavy scenarios from the INTERACTION dataset, contrasting with the 914 urban/highway scenarios of WOMD. As shown in Table 2, INTERACTION's complex interaction dynamics lead to higher error rates, particularly in collision and wrong-way incidents, across all models. In Phase I (zero-shot), CDPT achieves the best performance, with the lowest collision and log divergence, and competitive results in off-road, wrong-way, and kinematic infeasibility. The teacher model shows higher errors, especially in collision and wrong-way, reflecting limitations in decision-intensive scenarios. Ablation results show that FSD contributes the most significantly to overall performance gains, especially in reducing collisions and log divergences compared to the baseline teacher model. However, RSD and CSD are essential for controlling kinematic infeasibility and wrong-way rates, as their absence leads to higher values in these metrics. As illustrated in Figure 4, CDPT demonstrates improved compliance, stability, and safety in diverse interactive scenarios. These findings highlight CDPT's superior zero-shot generalization for generating safe, contextually appropriate behaviors in interactive scenarios.

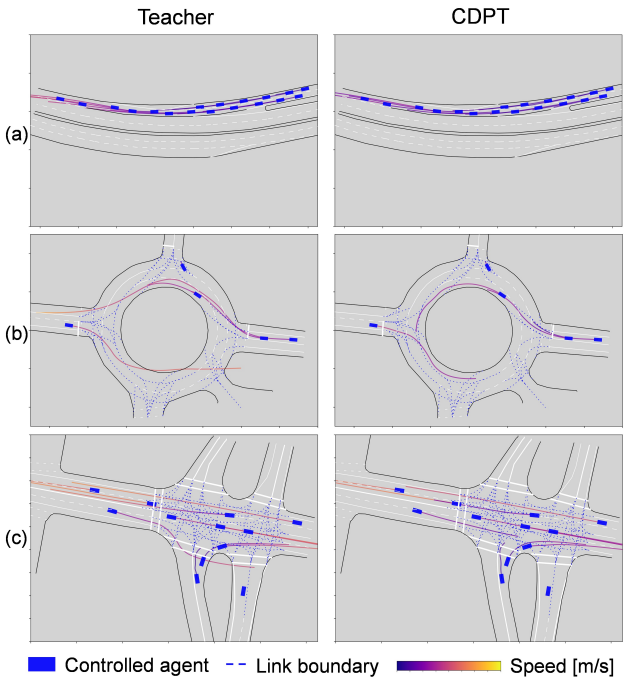

Figure 4: Illustrations of closed loop simulations in INTERACTION dataset: (a) In a merging scenario, CDPT better coordinates multiple agents, highlighting its capacity for interaction modeling. (b) In a roundabout, CDPT maintains lane stability, while the teacher exhibits frequent drifting and off-road behavior. (c) In a complex intersection scenario, CDPT significantly reduces road boundary violations and collisions compared to the teacher.

Continual Learning for Domain Adaptation

In Phase II, CDPT employs continual distillation, using few-shot samples from the target domain to generate synthetic scenarios, enabling adaptation to novel environments without large-scale retraining. As shown in Table 2, CDPT

FSD	RSD	CSD	ConD	WOMD					INTERACTION				
				C [%]↓	O [%]↓	W [%]↓	L [m]↓	K [%]↓	C [%]↓	O [%]↓	W [%]↓	L [m]↓	K [%]↓
Teacher				10.378	5.874	2.600	2.313	0.230	12.441	7.125	7.312	2.587	0.341
✓				8.311	5.785	2.339	1.919	0.329	10.207	6.832	7.154	2.354	0.412
	✓			12.118	6.539	**1.594**	2.398	**0.016**	13.541	7.325	6.112	2.654	**0.041**
		✓		9.475	5.765	2.444	2.309	0.151	11.507	6.387	6.987	2.541	0.141
✓			✓	8.151	5.822	2.384	1.998	0.272	8.154	6.125	6.421	2.187	0.487
	✓		✓	11.859	6.579	1.966	2.378	0.014	12.854	6.987	5.987	2.421	0.054
		✓	✓	9.198	5.603	2.251	2.349	0.194	8.252	4.027	6.154	2.321	0.378
✓		✓		8.336	5.694	2.345	1.928	0.172	9.205	6.031	6.741	2.312	0.212
✓	✓			8.300	5.623	2.409	2.340	0.141	9.743	5.906	6.654	2.361	0.481
✓		✓	✓	8.219	6.244	2.509	1.967	0.088	8.154	5.124	6.321	**2.104**	0.154
✓	✓		✓	8.286	6.037	2.511	2.277	0.131	7.738	3.578	6.112	2.154	0.964
✓	✓	✓		7.711	**5.191**	1.997	**1.854**	0.102	9.149	5.896	6.167	2.321	0.194
CDPT (full)				**7.644**	5.388	2.224	1.887	0.136	**6.853**	**3.117**	**5.914**	2.112	0.251

Table 2: Ablation study on the components of CDPT and performance comparison of methods in WOMD and INTERACTION datasets for closed-loop simulation, with phase I and phase II metrics: off-road (O), collision (C), wrong-way (W), kinematic infeasibility (K), and log divergences (L). Best values are **bolded**, and second-best values are underlined.

achieves the lowest error rates, including collision (6.853%) and log divergence (2.112 m) in INTERACTION, and maintains superior performance in WOMD. Compared to the teacher, which struggles in decision-intensive scenarios, and other ablation variants, CDPT consistently demonstrates better performance, with the continual distillation strategy further enhancing gains from FSD while preserving the controls on kinematic infeasibility and wrong-way metrics provided by RSD and CSD. These results underscore CDPT's robustness in adapting to diverse domains, generating realistic, interaction-aware behaviors critical for scalable autonomous driving testing.

Performance of CDPT in AV testing

Furthermore, we evaluated the effectiveness of the CDPT framework in closed-loop AV testing using the OnSite platform (OnSite 2025), a standardized benchmark for zero-shot scenario generation and safety-focused simulation. Each test case provides a high-definition map and a 31-frame prefix of agent states, requiring models to synthesize full traffic scenes based solely on partial observations. The benchmark includes a wide range of traffic scenarios such as merges, roundabouts, signalized intersections, and urban environments with vulnerable road users. This setting supports rigorous evaluation of generalization, realism, and causal reasoning without fine-tuning. The ego vehicle is evaluated under five different external planners, each tested separately on the same generated scenario. BVs are generated by CDPT or one of six baselines: a rule-based IDM controller (Dosovitskiy et al. 2017), a joint trajectory generator MJTG (Tian et al. 2025), a next-token prediction model SMART (Wu et al. 2024), two reinforcement learning methods SAC (Haarnoja et al. 2018) and EGPO (Fan et al. 2025b), and a diffusion-based generator VBD (Huang et al. 2024). The quality of the generated scenario is evaluated using five metrics: safety, consistency of distribution, comfort, effectiveness of the test, and total score. As shown in Table 3, CDPT consistently outperforms all baselines, particularly in test effectiveness and total score, demonstrating its ability to generate realistic and diverse traffic that improves planner robustness in safety-critical evaluations.

Method	SS ↑	KC ↑	CS ↑	TE ↑	TS ↑
IDM	16.73	5.84	5.77	6.06	34.40
SAC	16.11	4.59	7.04	9.26	37.00
MJTG	19.61	**13.08**	5.94	7.16	45.80
EGPO	26.03	5.31	**8.77**	13.71	53.82
VBD	**29.96**	7.47	7.62	16.18	61.23
SMART	22.88	8.60	6.88	13.24	51.60
CDPT	**29.96**	7.76	7.65	**17.92**	**63.29**

Table 3: Performance comparison of different methods in OnSite benchmark with Safety Score (SS), Kinetic distribution consistency (KC), Comfort Score (CS), Test Effectiveness (TE), and Total Score (TS). Best values are **bolded**, and second-best values are underlined.

Conclusion

In this paper, we propose a novel two-stage self-distillation framework, termed *Causal Driving Pattern Transfer (CDPT)*, which enhances both simulation fidelity and cross-domain generalization in multi-agent traffic simulation. CDPT leverages diffusion models and a structured knowledge distillation pipeline to effectively transfer causal behavioral patterns across diverse datasets. Extensive experiments demonstrate that CDPT produces realistic, interaction-aware traffic behaviors, significantly improving generalization and scalability for autonomous driving simulation testing. Future work will investigate extending this distillation framework to other simulation models, such as autoregressive or GPT architectures, to further enhance its applicability across diverse traffic simulation contexts.

Acknowledgements

This research is partially supported by the National Natural Science Foundation of China (Nos. 52125208, 52402374) and the Fundamental Research Funds for the Central Universities.

References

Caesar, H.; Bankiti, V.; Lang, A. H.; Vora, S.; Liong, V. E.; Xu, Q.; Krishnan, A.; Pan, Y.; Baldan, G.; and Beijbom, O. 2020. nuScenes: A multimodal dataset for autonomous driving. arXiv:1903.11027.

Dosovitskiy, A.; Ros, G.; Codevilla, F.; Lopez, A.; and Koltun, V. 2017. CARLA: An open urban driving simulator. In *Conference on robot learning*, 1–16. PMLR.

Fan, J.; Ni, Y.; Chen, Y.; Li, S.; Sun, J.; and Sun, J. 2025a. Interactive Adversarial Scenario Generation for Autonomous Driving: A Continual Learning Framework with Safety Constraints. In *IEEE International Conference on Intelligent Transportation Systems (ITSC)*. Gold Coast, AU.

Fan, J.; Ni, Y.; Yang, Y.; Zheng, W.; Sun, J.; and Sun, J. 2025b. Learning to Model Diverse Interactive Traffic with Driving Tendency-Guided Policy Optimization. *2025 IEEE Intelligent Vehicles Symposium (IV)*, 2047–2053.

Fan, J.; Ni, Y.; Zhao, D.; Hang, P.; and Sun, J. 2025c. Toward Proactive-Aware Autonomous Driving: A Reinforcement Learning Approach Utilizing Expert Priors During Unprotected Turns. *IEEE Transactions on Intelligent Transportation Systems*, 26(3): 3700–3712.

Gulino, C.; Fu, J.; Luo, W.; Tucker, G.; Bronstein, E.; Lu, Y.; Harb, J.; Pan, X.; Wang, Y.; Chen, X.; et al. 2023. Waymax: An accelerated, data-driven simulator for large-scale autonomous driving research. *Advances in Neural Information Processing Systems*, 36: 7730–7742.

Guo, Z.; Gao, X.; Zhou, J.; Cai, X.; and Shi, B. 2023. Scenedm: Scene-level multi-agent trajectory generation with consistent diffusion models. *arXiv preprint arXiv:2311.15736*.

Haarnoja, T.; Zhou, A.; Abbeel, P.; and Levine, S. 2018. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. arXiv:1801.01290.

Hinton, G.; Vinyals, O.; and Dean, J. 2015. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*.

Huang, Z.; Zhang, Z.; Vaidya, A.; Chen, Y.; Lv, C.; and Fisac, J. F. 2024. Versatile Behavior Diffusion for Generalized Traffic Agent Simulation. *arXiv preprint arXiv:2404.02524*.

Jiang, M.; Bai, Y.; Cornman, A.; Davis, C.; Huang, X.; Jeon, H.; Kulshrestha, S.; Lambert, J.; Li, S.; Zhou, X.; et al. 2024. Scenediffuser: Efficient and controllable driving simulation initialization and rollout. *Advances in Neural Information Processing Systems*, 37: 55729–55760.

Liu, J.; Xu, C.; Hang, P.; Sun, J.; Ding, M.; Zhan, W.; and Tomizuka, M. 2024. Language-driven policy distillation for cooperative driving in multi-agent reinforcement learning. *arXiv preprint arXiv:2410.24152*.

Montali, N.; Lambert, J.; Mougin, P.; Kuefler, A.; Rhinehart, N.; Li, M.; Gulino, C.; Emrich, T.; Yang, Z.; Whiteson, S.; et al. 2023. The waymo open sim agents challenge. *Advances in Neural Information Processing Systems*, 36: 59151–59171.

Monti, A.; Porrello, A.; Calderara, S.; Coscia, P.; Ballan, L.; and Cucchiara, R. 2022. How many observations are enough? knowledge distillation for trajectory forecasting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6553–6562.

Nichol, A. Q.; and Dhariwal, P. 2021. Improved denoising diffusion probabilistic models. In *International conference on machine learning*, 8162–8171. PMLR.

OnSite. 2025. Open Natural Driving Intelligence Automotive Simulation Test Environment. Available at: https://www.onsite.com.cn/#/dist/announcement3. Accessed: 21-Jun-2025.

Rong, G.; Shin, B. H.; Tabatabaee, H.; Lu, Q.; Lemke, S.; Možeiko, M.; Boise, E.; Uhm, G.; Gerow, M.; Mehta, S.; et al. 2020. LGSVL simulator: A high fidelity simulator for autonomous driving. In *2020 IEEE 23rd International conference on intelligent transportation systems (ITSC)*, 1–6. IEEE.

Song, J.; Ren, H.; Sadigh, D.; and Ermon, S. 2018. Multi-agent generative adversarial imitation learning. *arXiv preprint arXiv:1807.09936*.

Sun, J.; and Kim, J. 2024. Modelling two-dimensional driving behaviours at unsignalised intersection using multi-agent imitation learning. *Transportation Research Part C: Emerging Technologies*, 165: 104702.

Sun, J.; and Yang, H. 2024. Learning two-dimensional merging behaviour from vehicle trajectories with imitation learning. *Transportation Research Part C: Emerging Technologies*, 160: 104530.

Sun, J.; Zhang, H.; Zhou, H.; Yu, R.; and Tian, Y. 2022. Scenario-Based Test Automation for Highly Automated Vehicles: A Review and Paving the Way for Systematic Safety Assurance. *IEEE Transactions on Intelligent Transportation Systems*, 23(9): 14088–14103.

Tian, Y.; Zheng, W.; Shao, Y.; Zhang, H.; and Sun, J. 2025. MJTG: A Multi-vehicle Joint Trajectory Generator for Complex and Rare Scenarios. *IEEE Transactions on Vehicular Technology*, 1–14.

Treiber, M.; and Kesting, A. 2013. *Traffic flow dynamics*. Springer.

Wang, S.; Ni, Y.; Miao, C.; Sun, J.; and Sun, J. 2025. A multiagent social interaction model for autonomous vehicle testing. *Communications in Transportation Research*.

Wilson, B.; Qi, W.; Agarwal, T.; Lambert, J.; Singh, J.; Khandelwal, S.; Pan, B.; Kumar, R.; Hartnett, A.; Pontes, J. K.; et al. 2023. Argoverse 2: Next generation datasets for self-driving perception and forecasting. *arXiv preprint arXiv:2301.00493*.

Wu, W.; Feng, X.; Gao, Z.; and Kan, Y. 2024. SMART: scalable multi-agent real-time motion generation via next-token prediction. *Advances in Neural Information Processing Systems*, 37: 114048–114071.

Yan, X.; Zou, Z.; Feng, S.; Zhu, H.; Sun, H.; and Liu, H. X. 2023. Learning naturalistic driving environment with statistical realism. *Nature communications*, 14(1): 2037.

Yao, Y.; Goehring, D.; and Reichardt, J. 2025. Beyond In-Distribution Performance: A Cross-Dataset Study of Trajectory Prediction Robustness. *arXiv preprint arXiv:2501.15842*.

Zhan, W.; Sun, L.; Wang, D.; Shi, H.; Clausse, A.; Naumann, M.; Kummerle, J.; Konigshof, H.; Stiller, C.; de La Fortelle, A.; et al. 2019. Interaction dataset: An international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps. *arXiv preprint arXiv:1910.03088*.

Zhang, Z.; Karkus, P.; Igl, M.; Ding, W.; Chen, Y.; Ivanovic, B.; and Pavone, M. 2025. Closed-Loop Supervised Fine-Tuning of Tokenized Traffic Models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Zhou, Z.; Haibo, H.; Chen, X.; Wang, J.; Guan, N.; Wu, K.; Li, Y.-H.; Huang, Y.-K.; and Xue, C. J. 2024. Behaviorgpt: Smart agent simulation for autonomous driving with next-patch prediction. *Advances in Neural Information Processing Systems*, 37: 79597–79617.