Novelty Assessment Report

Paper: FlowAD: Ego-Scene Interactive Modeling for Autonomous Driving
PDF URL: https://openreview.net/pdf?id=m4JpoJRgAr
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-01

Abstract

Effective environment modeling is the foundation for autonomous driving, underpinning tasks from perception to planning. However, current paradigms often inadequately consider the feedback of ego motion to the observation, which leads to an incomplete understanding of the driving process and consequently limits the planning capability. To address this issue, we introduce a novel ego-scene interactive modeling paradigm. Inspired by human recognition, the paradigm represents ego-scene interaction as the scene flow relative to the ego-vehicle. This conceptualization allows for modeling ego-motion feedback within a feature learning pattern, advantageously utilizing existing log-replay datasets rather than relying on scenario simulations. We specifically propose FlowAD, a general flow-based framework for autonomous driving. Within it, an ego-guided scene partition first constructs basic flow units to quantify scene flow. The ego-vehicle's forward direction and steering velocity directly shape the partition, which reflects ego motion. Then, based on flow units, spatial and temporal flow predictions are performed to model dynamics of scene flow, encompassing both spatial displacement and temporal variation. The final task-aware enhancement exploits learned spatio-temporal flow dynamics to benefit diverse tasks through object and region-level strategies. We also propose a novel Frames before Correct Planning (FCP) metric to assess the scene understanding capability. Experiments in both open and closed-loop evaluations demonstrate FlowAD's generality and effectiveness across perception, end-to-end planning, and VLM analysis. Notably, FlowAD reduces 19\% collision rate over SparseDrive with FCP improvements of 1.39 frames (60\%) on nuScenes, and achieves an impressive driving score of 51.77 on Bench2Drive, proving the superiority. Code, model, and configurations will be released.

Core Task Landscape

This paper addresses: **Ego-Scene Interactive Modeling for Autonomous Driving**

A total of **50 papers** were analyzed and organized into a taxonomy with **32 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **World Models and Generative Scene Simulation**
- **End-to-End Planning and Decision-Making**
- **Trajectory Prediction and Motion Forecasting**
- **Perception and Scene Understanding**
- **Simulation and Testing Environments**
- **Multi-Agent Systems and Communication**
- **Specialized Interaction Scenarios**

Complete Taxonomy Tree

- Ego-Scene Interactive Modeling for Autonomous Driving Survey Taxonomy
- World Models and Generative Scene Simulation
  - Video-Based Generative World Models (5 papers)
  - [2] GAIA-1: A Generative World Model for Autonomous Driving (Hu, 2023) View paper
  - [9] OASim: an Open and Adaptive Simulator based on Neural Rendering for Autonomous Driving (Yan, 2024) View paper
  - [21] Other vehicle trajectories are also needed: A driving world model unifies ego-other vehicle trajectories in video latent space (Zhu Jian, 2025) View paper
  - [28] Generating Multimodal Driving Scenes via Next-Scene Prediction (WU Yanhao, 2025) View paper
  - [32] Geodrive: 3d geometry-informed driving world model with precise action control (Chen, 2025) View paper
  - Occupancy and Spatial World Models (3 papers)
  - [6] OccWorld: Learning a 3D Occupancy World Model for Autonomous Driving (Wenzhao Zheng, 2023) View paper
  - [13] Driving in the Occupancy World: Vision-Centric 4D Occupancy Forecasting and Planning via World Models for Autonomous Driving (Yu Yang, 2024) View paper
  - [36] COME: Adding Scene-Centric Forecasting Control to Occupancy World Model (Shi Yi-ning, 2025) View paper
  - Scene Representation Learning and Pre-training (2 papers)
  - [8] GASP: Unifying Geometric and Semantic Self-Supervised Pre-training for Autonomous Driving (Ljungbergh, 2025) View paper
  - [33] Driveworld: 4d pre-trained scene understanding via world models for autonomous driving (Chen Min, 2024) View paper
  - World Model Surveys and Taxonomies (1 papers)
  - [1] A survey of world models for autonomous driving (Feng Tuo, 2025) View paper
- End-to-End Planning and Decision-Making
  - Generative and Diffusion-Based Planning (3 papers)
  - [3] GenAD: Generative End-to-End Autonomous Driving (Zheng, 2024) View paper
  - [4] DiffusionDrive: Truncated Diffusion Model for End-to-End Autonomous Driving (Bencheng Liao, 2024) View paper
  - [41] DAP: A Discrete-token Autoregressive Planner for Autonomous Driving (Bowen Ye, 2025) View paper

- Deterministic and Reinforcement Learning-Based Planning (3 papers)
- [29] Cimrl: Combining imitation and reinforcement learning for safe autonomous driving (Rohanimanesh, 2024) View paper
- [30] Augmenting Reinforcement Learning With Transformer-Based Scene Representation Learning for Decision-Making of Autonomous Driving (Haochen Liu, 2022) View paper
- [47] Deep reinforcement and IL for autonomous driving: A review in the CARLA simulation environment (Piotr Czechowski, 2025) View paper
- Interaction-Aware and Graph-Based Planning (3 papers)
- [5] GraphAD: Interaction Scene Graph for End-to-end Autonomous Driving (Yunpeng Zhang, 2024) View paper
- [11] ICP-AD: Interactive-Cognition Planning for Autonomous Driving (Yichi Zhang, 2025) View paper
- [12] Reasoning Multi-Agent Behavioral Topology for Interactive Autonomous Driving (Li Chen, 2024) View paper
- Latent World Model-Enhanced Planning (2 papers)
- [42] Enhancing End-to-End Autonomous Driving with Latent World Model (Li, 2024) View paper
- [43] PIE: Perception and Interaction Enhanced End-to-End Motion Planning for Autonomous Driving (Yuan, 2025) View paper
- Vectorized and Structured Scene Planning (1 papers)
- [19] Vad: Vectorized scene representation for efficient autonomous driving (Bo Jiang, 2023) View paper
- Bidirectional and Future-Aware Planning (1 papers)
- [16] Future-aware end-to-end driving: Bidirectional modeling of trajectory planning and scene evolution (Zhang Bo-zhou, 2025) View paper
- Knowledge Distillation and Multi-Mode Planning (1 papers)
- [23] DistillDrive: End-to-End Multi-Mode Autonomous Driving Distillation by Isomorphic Hetero-Source Planning Model (Yu Rui, 2025) View paper
- Vision-Language Model-Guided Planning (1 papers)
- [17] VLMs Guided Interpretable Decision Making for Autonomous Driving (Xin Hu, 2025) View paper
- Ego-Scene Interactive Flow-Based Planning ★ (1 papers)
- [0] FlowAD: Ego-Scene Interactive Modeling for Autonomous Driving (Anon et al., 2026) View paper

- Trajectory Prediction and Motion Forecasting
  - Game-Theoretic and Interactive Prediction (1 papers)
  - [7] GameFormer: Game-theoretic Modeling and Learning of Transformer-based Interactive Prediction and Planning for Autonomous Driving (Zhiyu Huang, 2023) View paper
  - Scene-Consistent and Graph-Based Prediction (2 papers)
  - [34] Implicit latent variable model for scene-consistent motion forecasting (Sergio Casas, 2020) View paper
  - [35] Scene transformer: A unified architecture for predicting multiple agent trajectories (Ngiam, 2021) View paper
  - Motion Decoupling and Spatiotemporal Modeling (1 papers)
  - [14] DeMo++: Motion Decoupling for Autonomous Driving (Zhang Bo-zhou, 2025) View paper
  - Future Context-Aware Prediction (1 papers)
  - [44] FutureNet-LOF: Joint trajectory prediction and lane occupancy field prediction with future context encoding (Mingkun Wang, 2025) View paper
  - Pedestrian Trajectory Prediction (1 papers)
  - [27] Pedestrian trajectory prediction in pedestrian-vehicle mixed environments: A systematic review (Ghafurian, 2023) View paper
  - Ego Trajectory Prediction in BEV (1 papers)
  - [31] BEVSeg2GTA: Joint Vehicle Segmentation and Graph Neural Networks for Ego Vehicle Trajectory Prediction in Bird's-Eye-View (Sharma Sushil, 2024) View paper
- Perception and Scene Understanding
  - Egocentric Spatial-Temporal Interaction Modeling (1 papers)
  - [50] Learning 3d-aware egocentric spatial-temporal interaction via graph convolutional networks (Chengxi Li, 2020) View paper
  - Generalized Scene Understanding (1 papers)
  - [39] PreGSU: A Generalized Traffic Scene Understanding Model for Autonomous Driving Based on Pretrained Graph Attention Network (Wang Yu-ning, 2025) View paper
  - Pedestrian and Human Interaction Perception (1 papers)
  - [46] Pedestrian Facial Attention Detection Using Deep Fusion and Multi-Modal Fusion Classifier (Jing Lian, 2025) View paper
  - Egocentric Video Captioning (1 papers)
  - [20] Egocentric vehicle dense video captioning (Feiyu Chen, 2024) View paper
- Simulation and Testing Environments
  - Data-Driven and Replay-Based Simulators (1 papers)
  - [22] Waymax: An accelerated, data-driven simulator for large-scale autonomous driving research (Gulino, 2023) View paper
  - Neural Rendering and Adaptive Simulation (1 papers)
  - [48] Vid2sim: Realistic and interactive simulation from video for urban navigation (Ziyang Xie, 2025) View paper
  - Interactive Traffic Simulation (2 papers)
  - [10] Towards interactive autonomous vehicle testing: Vehicle-under-test-centered traffic simulation (Yiru Liu, 2024) View paper
  - [49] InterSim: Interactive Traffic Simulation via Explicit Relation Modeling (Qiao Sun, 2022) View paper
  - Scenario Generation and Augmentation (2 papers)
  - [15] Learning naturalistic driving environment with statistical realism (Xintao Yan, 2023) View paper
  - [18] UniGen: Unified Modeling of Initial Agent States and Trajectories for Generating Autonomous Driving Scenarios (Reza Mahjourian, 2024) View paper
  - Competitive and Multi-Agent Behavior Modeling (2 papers)
  - [37] Learning to Model Diverse Driving Behaviors in Highly Interactive Autonomous Driving Scenarios With Multiagent Reinforcement Learning (Weiwei Liu, 2024) View paper
  - [38] Modelling competitive behaviors in autonomous driving under generative world model (Guanren Qiao, 2024) View paper
- Multi-Agent Systems and Communication
  - Communicative World Models (1 papers)
  - [24] Ego-centric Learning of Communicative World Models for Autonomous Driving (Wang Hang, 2025) View paper
  - LLM-Based Multi-Agent Systems (2 papers)
  - [25] Multi-agent autonomous driving systems with large language models: A survey of recent advances (Wu, 2025) View paper

    - [26] Multi-agent autonomous driving systems with large language models: A survey of recent advances, resources, and future directions (Yaozu Wu, 2025) View paper
    - Interactive Autonomous Navigation (1 papers)
    - [40] Interactive autonomous navigation with internal state inference and interactivity estimation (Jiachen Li, 2024) View paper
- Specialized Interaction Scenarios (1 papers)
    - [45] Autonomous Vehicle and Pedestrian Interaction (Elaa Elgharbi, 2024) View paper

Narrative

Core task: ego-scene interactive modeling for autonomous driving. This field addresses how an autonomous vehicle perceives, predicts, and plans within dynamic traffic environments where the ego agent must continuously reason about interactions with other road users. The taxonomy organizes research into several major branches: World Models and Generative Scene Simulation focuses on learning predictive models of future scene evolution, often using generative architectures like diffusion or autoregressive frameworks (e.g., GAIA-1[2], GenAD[3], OccWorld[6]); End-to-End Planning and Decision-Making emphasizes direct mapping from perception to control, including flow-based and reinforcement learning approaches; Trajectory Prediction and Motion Forecasting targets anticipating the future paths of surrounding agents through joint or marginal modeling (e.g., GameFormer[7], Scene Transformer[35]); Perception and Scene Understanding covers representation learning and semantic reasoning; Simulation and Testing Environments provides platforms for closed-loop evaluation (e.g., Waymax[22]); Multi-Agent Systems and Communication explores coordination and information exchange; and Specialized Interaction Scenarios examines context-specific behaviors such as pedestrian crossings or merging maneuvers.

A central tension across these branches is the trade-off between interpretability and end-to-end performance: world models offer explicit future rollouts and controllable generation but may struggle with real-time constraints, while end-to-end planners can be faster yet less transparent. Within the End-to-End Planning branch, FlowAD[0] adopts a flow-based formulation for interactive planning, distinguishing itself from diffusion-based methods like DiffusionDrive[4] by leveraging normalizing flows for tractable likelihood modeling and efficient sampling. This approach contrasts with graph-structured planners such as GraphAD[5], which explicitly encode spatial relationships, and with reinforcement learning pipelines that require extensive online interaction. FlowAD[0] thus occupies a niche emphasizing probabilistic modeling with closed-form density estimation, bridging generative scene simulation and direct planning in a computationally efficient manner while maintaining a degree of interpretability through its flow architecture.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

Sibling Subtopics

- **Bidirectional and Future-Aware Planning** (leaves: 1, papers: 1)
- Scope: Planners modeling bidirectional dependencies between ego trajectory and scene evolution, explicitly encoding future context.
- Exclude: Excludes one-shot planning without future feedback; see Generative and Diffusion-Based Planning or Deterministic and Reinforcement Learning-Based Planning.
- **Deterministic and Reinforcement Learning-Based Planning** (leaves: 1, papers: 3)
- Scope: Methods employing deterministic encoders, RL policies, or hybrid imitation-RL frameworks for trajectory planning and control.
- Exclude: Excludes generative multi-mode planning; see Generative and Diffusion-Based Planning.
- **Generative and Diffusion-Based Planning** (leaves: 1, papers: 3)
- Scope: End-to-end planners using generative models or diffusion processes to produce multi-mode trajectory distributions conditioned on scene context.
- Exclude: Excludes deterministic or single-mode planning; see Deterministic and Reinforcement Learning-Based Planning.
- **Interaction-Aware and Graph-Based Planning** (leaves: 1, papers: 3)
- Scope: Planners explicitly modeling agent interactions via graph structures or relational reasoning to inform trajectory decisions.
- Exclude: Excludes methods without explicit interaction modeling; see Generative and Diffusion-Based Planning or Deterministic and Reinforcement Learning-Based Planning.
- **Knowledge Distillation and Multi-Mode Planning** (leaves: 1, papers: 1)
- Scope: Approaches using knowledge distillation from heterogeneous planning models to learn multi-mode driving behaviors.
- Exclude: Excludes single-source or non-distillation methods; see Generative and Diffusion-Based Planning.
- **Latent World Model-Enhanced Planning** (leaves: 1, papers: 2)
- Scope: End-to-end planners augmented with latent world models for self-supervised feature learning or future state prediction to improve decision-making.
- Exclude: Excludes explicit generative scene simulation; see World Models and Generative Scene Simulation.
- **Vectorized and Structured Scene Planning** (leaves: 1, papers: 1)
- Scope: Methods representing scenes as vectorized agents and map elements for efficient planning, avoiding dense rasterization.
- Exclude: Excludes dense occupancy or pixel-based representations; see Occupancy and Spatial World Models.
- **Vision-Language Model-Guided Planning** (leaves: 1, papers: 1)
- Scope: Methods leveraging vision-language models for interpretable high-level decision-making via visual question answering or reasoning.
- Exclude: Excludes purely vision-based or non-VLM methods; see other subcategories.

Contributions Analysis

Overall novelty summary. The paper introduces FlowAD, a flow-based framework for autonomous driving that models ego-scene interaction through scene flow relative to the ego vehicle. Within the taxonomy, it occupies a unique leaf node under End-to-End Planning and Decision-Making labeled 'Ego-Scene Interactive Flow-Based Planning,' with no sibling papers in this specific category. This positioning suggests the work addresses a relatively sparse research direction, as the taxonomy contains 50 papers across approximately 36 topics, yet this particular formulation of ego-motion feedback through scene flow appears underexplored in the surveyed literature.

The taxonomy reveals that FlowAD sits within a broader End-to-End Planning branch containing eight subcategories, including Generative and Diffusion-Based Planning (e.g., DiffusionDrive), Deterministic and Reinforcement Learning-Based Planning, and Interaction-Aware and Graph-Based Planning (e.g., GraphAD). Neighboring branches include World Models and Generative Scene Simulation, which focuses on explicit future state prediction through occupancy grids or video generation, and Trajectory Prediction and Motion Forecasting, which emphasizes multi-agent dynamics without integrated planning. FlowAD's flow-based formulation distinguishes it from diffusion methods by offering tractable likelihood modeling, while its ego-guided scene partition diverges from graph-structured approaches by directly encoding ego motion into spatial decomposition.

Among 29 candidates examined across three contributions, no clearly refutable prior work was identified. The ego-scene interactive modeling paradigm examined 9 candidates with 0 refutations, the FlowAD framework examined 10 candidates with 0 refutations, and the

Frames before Correct Planning metric examined 10 candidates with 0 refutations. This suggests that within the limited search scope—primarily top-K semantic matches and citation expansion—the specific combination of scene flow representation, ego-motion feedback modeling, and flow-based planning appears novel. However, the analysis explicitly notes this is not an exhaustive literature search, and the absence of refutations reflects the examined sample rather than comprehensive field coverage.

Given the limited search scope of 29 candidates, the work appears to introduce a distinctive approach within end-to-end planning by formalizing ego-scene interaction as relative scene flow and leveraging normalizing flows for probabilistic trajectory generation. The sparse population of its taxonomy leaf and the absence of refutable candidates among examined papers suggest potential novelty, though the analysis cannot rule out relevant prior work outside the semantic search radius or in adjacent subfields not fully captured by the taxonomy structure.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Ego-scene interactive modeling paradigm for autonomous driving

Description: The authors propose a new paradigm that models the feedback of ego motion to environmental observation by representing ego-scene interaction as scene flow relative to the ego-vehicle. This approach enables modeling ego-motion feedback within feature learning using log-replay datasets rather than requiring scenario simulations.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Scene Flow Specifications: Encoding and Monitoring Rich Temporal Safety Properties of Autonomous Systems

URL: View paper

Brief Assessment

Scene Flow Specifications[60] focuses on runtime monitoring and specification languages for safety properties using symbolic entities across time, not on modeling ego-motion feedback within feature learning for autonomous driving perception and planning tasks.

---

2. PillarFlowNet: A Real-time Deep Multitask Network for LiDAR-based 3D Object Detection and Scene Flow Estimation

URL: View paper

Brief Assessment

PillarFlowNet[64] focuses on LiDAR-based scene flow estimation for object detection and motion understanding, not on modeling ego-motion feedback within feature learning. The candidate addresses scene flow as a perception output for understanding other agents' motion, while the original paper proposes using scene flow relative to ego-vehicle as a mechanism to model how ego-motion affects environmental observation within the learning process.

---

3. Using Sceneâ Flow to Improve Predictions of Road Users in Motion With Respect to an Egoâ Vehicle

URL: View paper

Brief Assessment

Scene Flow Predictions[53] focuses on motion status classification of neighboring vehicles using scene-flow for tracking, not on modeling ego-motion feedback within feature learning for general autonomous driving tasks.

---

4. EgoFlowNet: Non-Rigid Scene Flow from Point Clouds with Ego-Motion Support

URL: View paper

Brief Assessment

EgoFlowNet[66] focuses on point-level scene flow estimation from LiDAR point clouds for non-rigid motion, not on modeling ego-scene interaction for autonomous driving planning using multi-view camera images and log-replay datasets.

---

5. SSF-MOS: Semantic Scene Flow Assisted Moving Object Segmentation for Autonomous Vehicles

URL: View paper

Brief Assessment

SSF-MOS[56] focuses on moving object segmentation using semantic scene flow in point clouds for detecting moving vs. static objects, not on modeling ego-motion feedback within feature learning for autonomous driving tasks as proposed in the original paper.

---

6. Active scene flow estimation for autonomous driving via real-time scene prediction and optimal decision

URL: View paper

Brief Assessment

Active Scene Flow[63] focuses on active scene flow estimation through changing ego-vehicle trajectory and sensor positions for better point cloud registration, not on modeling ego-motion feedback within feature learning for general autonomous driving tasks as in the original paper.

---

7. SSF-PAN: Semantic Scene Flow-Based Perception for Autonomous Navigation in Traffic Scenarios

URL: View paper

Brief Assessment

SSF-PAN[61] focuses on semantic scene flow for segmenting static vs. dynamic objects in LiDAR point clouds for SLAM and navigation. The original paper models ego-scene interaction through scene flow relative to the ego-vehicle within feature learning for end-to-end planning tasks, representing a fundamentally different approach and application domain.

---

8. Gaussianad: Gaussian-centric end-to-end autonomous driving

URL: View paper

Brief Assessment

Gaussianad[62] focuses on 3D Gaussian representations for scene modeling and predicts 3D flows for Gaussians with dynamic semantics, rather than modeling ego-scene interaction as scene flow relative to the ego-vehicle within a feature learning pattern using log-replay datasets. The technical approaches and objectives differ fundamentally.

---

9. Neural Eulerian Scene Flow Fields

URL: View paper

Brief Assessment

Neural Eulerian[65] focuses on scene flow estimation as a continuous space-time ODE for general 3D reconstruction across multiple domains (autonomous driving, tabletop scenes, tracking birds/tennis balls). The original paper's contribution is specifically about modeling ego-vehicle motion feedback within autonomous driving feature learning using log-replay datasets, which is a different architectural paradigm and application context.

Contribution 2: FlowAD: a general flow-based framework for autonomous driving

Description: The authors introduce FlowAD, a framework comprising three core components: ego-guided scene partition that constructs flow units reflecting ego motion, spatial and temporal flow predictions that model scene flow dynamics, and task-aware enhancement that exploits learned dynamics to benefit diverse downstream tasks.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. 20.6 LSPU: A Fully Integrated Real-Time LiDAR-SLAM SoC with Point-Neural-Network Segmentation and Multi-Level kNN Acceleration

URL: View paper

Brief Assessment

LSPU[55] focuses on hardware acceleration for LiDAR-SLAM systems with point neural networks, not on flow-based frameworks with ego-guided scene partition for autonomous driving perception and planning.

2. Ego-centric Learning of Communicative World Models for Autonomous Driving

URL: View paper

Brief Assessment

Communicative World Models[24] focuses on multi-agent reinforcement learning with lightweight communication between agents using world model latent representations, not on ego-guided scene partition or flow-based frameworks for single-agent autonomous driving tasks.

3. Using Sceneâ Flow to Improve Predictions of Road Users in Motion With Respect to an Egoâ Vehicle

URL: View paper

Brief Assessment

Scene Flow Predictions[53] presents a specific vision-based tracking approach (MOVE) for binary motion classification, not a general framework with ego-guided scene partition and task-aware enhancement for diverse downstream tasks.

4. SSF-MOS: Semantic Scene Flow Assisted Moving Object Segmentation for Autonomous Vehicles

URL: View paper

Brief Assessment

SSF-MOS[56] presents a unified framework for moving object segmentation with semantic information and ego-motion estimation, but does not propose a general flow-based framework with ego-guided scene partition, spatial/temporal flow predictions, and task-aware enhancement for diverse autonomous driving tasks.

5. Motion inspired unsupervised perception and prediction in autonomous driving

URL: View paper

Brief Assessment

Motion Inspired[51] focuses on unsupervised perception and prediction through motion-inspired scene decomposition and local flow estimation, rather than proposing a comprehensive flow-based framework with ego-guided scene partition, spatial/temporal flow predictions, and task-aware enhancement for multiple downstream tasks as in FlowAD.

6. Scalable scene flow from point clouds in the real world

URL: View paper

Brief Assessment

Scalable Scene Flow[58] focuses on scene flow estimation from point clouds using lidar data for 3D motion estimation, not on a general flow-based framework with ego-guided scene partition for autonomous driving tasks like perception, planning, and VLM analysis as proposed in the original paper.

7. Object scene flow for autonomous vehicles

URL: View paper

Brief Assessment

Object Scene Flow[54] focuses on 3D scene flow estimation for autonomous vehicles using rigid object decomposition and CAD model fitting, not on ego-guided scene partition or task-aware enhancement for diverse downstream tasks in autonomous driving frameworks.

8. Towards optical flow ego-motion compensation for moving object segmentation

URL: View paper

Brief Assessment

Optical Flow Compensation[59] focuses on ego-motion compensation for optical flow in robot navigation contexts, not on autonomous driving frameworks with ego-guided scene partition, spatial/temporal flow predictions, or task-aware enhancement for perception and planning tasks.

9. Joint Scene Flow Estimation and Moving Object Segmentation on Rotational LiDAR Data

URL: View paper

Brief Assessment

Joint Scene Flow[52] focuses on LiDAR-based scene flow estimation and moving object segmentation for point clouds, while FlowAD addresses ego-scene interactive modeling using multi-view camera images with ego-guided scene partition. The technical approaches, sensor modalities, and problem formulations are fundamentally different.

10. DiFSD: Ego-Centric Fully Sparse Paradigm with Uncertainty Denoising and Iterative Refinement for Efficient End-to-End Self-Driving

URL: View paper

Brief Assessment

DiFSD[57] focuses on an ego-centric fully sparse paradigm with sparse perception, hierarchical interaction, and iterative motion planning. It does not propose a flow-based framework with ego-guided scene partition or scene flow modeling as described in the original paper.

Contribution 3: Frames before Correct Planning (FCP) metric

Description: The authors introduce a new evaluation metric that quantifies the number of frames elapsed until a planner initiates a rational action in response to a given command, providing a statistical measure of a planner's comprehension of the driving process with ego-scene interaction.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. An efficient approach of time-optimal trajectory generation for the fully autonomous navigation of the quadrotor

URL: View paper

Brief Assessment

Time-Optimal Quadrotor[71] focuses on time-optimal trajectory generation for quadrotor navigation using NURBS curves and linear programming, not on evaluation metrics for autonomous driving planners. The paper addresses trajectory optimization for aerial vehicles rather than measuring planning response times in ground-based autonomous driving systems.

2. Continuous advantage learning for minimum-time trajectory planning of autonomous vehicles

URL: View paper

Brief Assessment

Continuous Advantage Learning[68] focuses on minimum-time trajectory planning using continuous advantage learning for autonomous vehicles, not on evaluation metrics for planning comprehension. No content related to FCP or similar metrics was found in the available context.

3. Enhancing unmanned vehicle navigation safety: real-time visual mapping with CNNs with optimized Bezier trajectory smoothing

URL: View paper

Brief Assessment

Visual Mapping CNNs[70] focuses on pothole detection and path planning optimization using CNNs with Bezier trajectory smoothing. It does not propose or discuss any metrics for measuring frames until correct planning in autonomous driving systems.

4. Predefined Time Trajectory Tracking Control of Underactuated Autonomous Underwater Vehicle*

URL: View paper

Brief Assessment

Predefined Time Tracking[75] focuses on trajectory tracking control for underwater vehicles with predefined time convergence. It does not address autonomous driving planning metrics or evaluation of planner comprehension capabilities.

5. Time-based trajectory control for an urban autonomous vehicle

URL: View paper

Brief Assessment

Time-based Trajectory Control[72] focuses on time-constrained trajectory control for autonomous vehicles in urban settings, examining time deviations and speed regulation. This is fundamentally different from the FCP metric, which statistically measures the number of frames until a planner initiates a rational action in response to commands.

6. Omnidirectional Autonomous Aggressive Perching of Unmanned Aerial Vehicle using Reinforcement Learning Trajectory Generation and Control

URL: View paper

Brief Assessment

Aggressive Perching[76] focuses on quadrotor trajectory generation and perching control using reinforcement learning, with no discussion of autonomous driving metrics or measuring frames until correct planning responses to commands.

7. An Approach to Generating Scenarios for Autonomous

Vehicles

URL: View paper

Brief Assessment

Generating Scenarios[73] focuses on generating virtual test scenarios for autonomous vehicles using game-theoretic trajectory generation, not on evaluation metrics for planning systems. The candidate does not propose or discuss metrics for measuring frames until correct planning.

8. Autonomous Reactive Masonry Construction using Collaborative Heterogeneous Aerial Robots with Experimental Demonstration

URL: View paper

Brief Assessment

Reactive Masonry Construction[69] focuses on autonomous aerial robotic construction using UAVs for masonry tasks. It does not address autonomous driving metrics or planning evaluation frameworks, making comparison with the FCP metric infeasible.

9. Time-Optimal Trajectory Planning and Tracking for Autonomous Vehicles

URL: View paper

Brief Assessment

Time-Optimal Planning[74] focuses on trajectory optimization and tracking for high-speed racing scenarios using hierarchical control (TRO and NMPC modules). It does not propose any evaluation metrics for measuring planning response time or frames until correct action initiation in autonomous driving contexts.

10. A real-time motion planner with trajectory optimization for autonomous vehicles

URL: View paper

Brief Assessment

Real-time Motion Planner[67] focuses on trajectory optimization for motion planning with cost functions and iterative optimization. It does not propose or discuss any evaluation metrics related to measuring frames until correct planning responses to commands.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] FlowAD: Ego-Scene Interactive Modeling for Autonomous Driving View paper
- [1] A survey of world models for autonomous driving View paper
- [2] GAIA-1: A Generative World Model for Autonomous Driving View paper
- [3] GenAD: Generative End-to-End Autonomous Driving View paper
- [4] DiffusionDrive: Truncated Diffusion Model for End-to-End Autonomous Driving View paper
- [5] GraphAD: Interaction Scene Graph for End-to-end Autonomous Driving View paper
- [6] OccWorld: Learning a 3D Occupancy World Model for Autonomous Driving View paper
- [7] GameFormer: Game-theoretic Modeling and Learning of Transformer-based Interactive Prediction and Planning for Autonomous Driving View paper
- [8] GASP: Unifying Geometric and Semantic Self-Supervised Pre-training for Autonomous Driving View paper
- [9] OASim: an Open and Adaptive Simulator based on Neural Rendering for Autonomous Driving View paper
- [10] Towards interactive autonomous vehicle testing: Vehicle-under-test-centered traffic simulation View paper
- [11] ICP-AD: Interactive-Cognition Planning for Autonomous Driving View paper
- [12] Reasoning Multi-Agent Behavioral Topology for Interactive Autonomous Driving View paper
- [13] Driving in the Occupancy World: Vision-Centric 4D Occupancy Forecasting and Planning via World Models for Autonomous Driving View paper
- [14] DeMo++: Motion Decoupling for Autonomous Driving View paper
- [15] Learning naturalistic driving environment with statistical realism View paper
- [16] Future-aware end-to-end driving: Bidirectional modeling of trajectory planning and scene evolution View paper
- [17] VLMs Guided Interpretable Decision Making for Autonomous Driving View paper
- [18] UniGen: Unified Modeling of Initial Agent States and Trajectories for Generating Autonomous Driving Scenarios View paper
- [19] Vad: Vectorized scene representation for efficient autonomous driving View paper
- [20] Egocentric vehicle dense video captioning View paper
- [21] Other vehicle trajectories are also needed: A driving world model unifies ego-other vehicle trajectories in video latent space View paper
- [22] Waymax: An accelerated, data-driven simulator for large-scale autonomous driving research View paper
- [23] DistillDrive: End-to-End Multi-Mode Autonomous Driving Distillation by Isomorphic Hetero-Source Planning Model View paper
- [24] Ego-centric Learning of Communicative World Models for Autonomous Driving View paper
- [25] Multi-agent autonomous driving systems with large language models: A survey of recent advances View paper
- [26] Multi-agent autonomous driving systems with large language models: A survey of recent advances, resources, and future directions View paper
- [27] Pedestrian trajectory prediction in pedestrian-vehicle mixed environments: A systematic review View paper
- [28] Generating Multimodal Driving Scenes via Next-Scene Prediction View paper
- [29] Cimrl: Combining imitation and reinforcement learning for safe autonomous driving View paper
- [30] Augmenting Reinforcement Learning With Transformer-Based Scene Representation Learning for Decision-Making of Autonomous Driving View paper
- [31] BEVSeg2GTA: Joint Vehicle Segmentation and Graph Neural Networks for Ego Vehicle Trajectory Prediction in Bird's-Eye-View View paper
- [32] Geodrive: 3d geometry-informed driving world model with precise action control View paper
- [33] Driveworld: 4d pre-trained scene understanding via world models for autonomous driving View paper
- [34] Implicit latent variable model for scene-consistent motion forecasting View paper
- [35] Scene transformer: A unified architecture for predicting multiple agent trajectories View paper
- [36] COME: Adding Scene-Centric Forecasting Control to Occupancy World Model View paper
- [37] Learning to Model Diverse Driving Behaviors in Highly Interactive Autonomous Driving Scenarios With Multiagent Reinforcement Learning View paper
- [38] Modelling competitive behaviors in autonomous driving under generative world model View paper
- [39] PreGSU: A Generalized Traffic Scene Understanding Model for Autonomous Driving Based on Pretrained Graph Attention Network View paper
- [40] Interactive autonomous navigation with internal state inference and interactivity estimation View paper
- [41] DAP: A Discrete-token Autoregressive Planner for Autonomous Driving View paper
- [42] Enhancing End-to-End Autonomous Driving with Latent World Model View paper
- [43] PIE: Perception and Interaction Enhanced End-to-End Motion Planning for Autonomous Driving View paper
- [44] FutureNet-LOF: Joint trajectory prediction and lane occupancy field prediction with future context encoding View paper
- [45] Autonomous Vehicle and Pedestrian Interaction View paper
- [46] Pedestrian Facial Attention Detection Using Deep Fusion and Multi-Modal Fusion Classifier View paper
- [47] Deep reinforcement and IL for autonomous driving: A review in the CARLA simulation environment View paper
- [48] Vid2sim: Realistic and interactive simulation from video for urban navigation View paper
- [49] InterSim: Interactive Traffic Simulation via Explicit Relation Modeling View paper
- [50] Learning 3d-aware egocentric spatial-temporal interaction via graph convolutional networks View paper
- [51] Motion inspired unsupervised perception and prediction in autonomous driving View paper
- [52] Joint Scene Flow Estimation and Moving Object Segmentation on Rotational LiDAR Data View paper
- [53] Using Sceneâ□□Flow to Improve Predictions of Road Users in Motion With Respect to an Egoâ□□Vehicle View paper
- [54] Object scene flow for autonomous vehicles View paper
- [55] 20.6 LSPU: A Fully Integrated Real-Time LiDAR-SLAM SoC with Point-Neural-Network Segmentation and Multi-Level kNN Acceleration View paper

- [56] SSF-MOS: Semantic Scene Flow Assisted Moving Object Segmentation for Autonomous Vehicles View paper
- [57] DiFSD: Ego-Centric Fully Sparse Paradigm with Uncertainty Denoising and Iterative Refinement for Efficient End-to-End Self-Driving View paper
- [58] Scalable scene flow from point clouds in the real world View paper
- [59] Towards optical flow ego-motion compensation for moving object segmentation View paper
- [60] Scene Flow Specifications: Encoding and Monitoring Rich Temporal Safety Properties of Autonomous Systems View paper
- [61] SSF-PAN: Semantic Scene Flow-Based Perception for Autonomous Navigation in Traffic Scenarios View paper
- [62] Gaussianad: Gaussian-centric end-to-end autonomous driving View paper
- [63] Active scene flow estimation for autonomous driving via real-time scene prediction and optimal decision View paper
- [64] PillarFlowNet: A Real-time Deep Multitask Network for LiDAR-based 3D Object Detection and Scene Flow Estimation View paper
- [65] Neural Eulerian Scene Flow Fields View paper
- [66] EgoFlowNet: Non-Rigid Scene Flow from Point Clouds with Ego-Motion Support View paper
- [67] A real-time motion planner with trajectory optimization for autonomous vehicles View paper
- [68] Continuous advantage learning for minimum-time trajectory planning of autonomous vehicles View paper
- [69] Autonomous Reactive Masonry Construction using Collaborative Heterogeneous Aerial Robots with Experimental Demonstration View paper
- [70] Enhancing unmanned vehicle navigation safety: real-time visual mapping with CNNs with optimized Bezier trajectory smoothing View paper
- [71] An efficient approach of time-optimal trajectory generation for the fully autonomous navigation of the quadrotor View paper
- [72] Time-based trajectory control for an urban autonomous vehicle View paper
- [73] An Approach to Generating Scenarios for Autonomous Vehicles View paper
- [74] Time-Optimal Trajectory Planning and Tracking for Autonomous Vehicles View paper
- [75] Predefined Time Trajectory Tracking Control of Underactuated Autonomous Underwater Vehicle* View paper
- [76] Omnidirectional Autonomous Aggressive Perching of Unmanned Aerial Vehicle using Reinforcement Learning Trajectory Generation and Control View paper