

Scaling Motion Prediction for Autonomous Driving: From Large-Scale Datasets to Modern Learning Paradigms

Mr. DANDU JAYABHARATH REDDY

Assistant Professor, Bachelor of Science in Information Technology, Sambhram University, Jizzax, Uzbekistan,
Email ID: bharath55.edu@gmail.com

NURMUXAMMADOV ANVAR

BS-IT Student, Sambhram University, Jizzax, Uzbekistan.

FARXODOV DILSHOD

BS-IT Student, Sambhram University, Jizzax, Uzbekistan.

Abstract

The rapid advancement of autonomous driving systems has been significantly driven by the availability of large-scale datasets. Early datasets primarily focused on perception tasks, while motion prediction and planning remained limited due to insufficient data and complexity of behavioral modeling. In this paper, we revisit the paradigm of large-scale motion prediction datasets and extend it with recent developments from 2021–2025. We present an updated perspective on dataset design, incorporating advances in transformer-based architectures, diffusion models, and closed-loop evaluation frameworks. Furthermore, we analyze how modern datasets and learning strategies improve trajectory forecasting and planning performance. Our findings highlight that scaling both data and model complexity remains critical for achieving robust, real-world autonomous driving systems.

Keywords: Autonomous driving, motion prediction, trajectory forecasting, transformers, diffusion models, self-driving datasets, multi-agent systems, motion planning, closed-loop simulation

1. Introduction

Large-scale datasets have played a crucial role in advancing machine learning, particularly in autonomous driving. Early efforts focused on perception tasks such as object detection and segmentation, enabled by datasets like KITTI and Waymo Open Dataset. However, perception alone is insufficient for safe autonomous navigation.

Modern autonomous driving pipelines require accurate **motion prediction**, where the future trajectories of surrounding agents (vehicles, pedestrians, cyclists) are estimated. This task is inherently complex due to multimodal behavior, uncertainty, and interactions between agents.

Earlier datasets addressed this limitation by introducing large-scale motion prediction data collected from fleets of self-driving vehicles. However, recent years have seen major shifts in both **data scale** and **modeling approaches**, including:

- Transformer-based architectures for sequential prediction
- Multi-agent interaction modeling
- Diffusion-based trajectory generation
- Closed-loop simulation benchmarks

This paper revisits the concept of large-scale motion prediction datasets and integrates recent advances to provide a modernized framework for autonomous driving research.

2. Related Work

2.1 Perception Datasets

Initial datasets such as KITTI, nuScenes, and Waymo Open Dataset enabled breakthroughs in 3D object detection and scene understanding. These datasets focus on processing raw sensor inputs such as LiDAR and camera data.

2.2 Motion Prediction Datasets

Motion prediction datasets evolved to include behavioral data and semantic maps. Notable improvements include:

- **Argoverse 2:** Expanded scale and richer annotations

- **Waymo Motion Dataset:** High-quality trajectory annotations with diverse scenarios
- **nuPlan:** Introduced planning-focused evaluation

Compared to earlier datasets, modern datasets emphasize:

- Long temporal horizons
- Rich map priors
- Multi-agent interactions

2.3 Learning-Based Motion Prediction

Traditional approaches relied on recurrent neural networks and convolutional architectures. Recent advances include:

- **Transformers:** Capture long-range dependencies and interactions
- **Graph Neural Networks (GNNs):** Model agent relationships
- **Diffusion Models:** Generate diverse and probabilistic trajectories

These methods significantly outperform earlier CNN-based baselines.

3. Dataset Design for Motion Prediction

A modern motion prediction dataset consists of three key components:

3.1 Scene Data

Scenes capture temporal sequences of traffic environments, typically including:

- Positions and velocities of agents
- Agent types (vehicle, pedestrian, cyclist)
- Temporal sequences (e.g., 20–30 seconds)

Modern datasets now include:

- Higher sampling rates
- Diverse environmental conditions
- Multi-city coverage

3.2 High-Definition Semantic Maps

Semantic maps provide contextual priors such as:

- Lane geometry and connectivity
- Traffic rules and signals
- Road structure and constraints

Recent improvements include vectorized map representations, enabling efficient learning and integration with neural networks.

3.3 Aerial and Rasterized Representations

Bird's-eye-view (BEV) representations remain standard for model input. However, newer approaches integrate:

- Vectorized representations (e.g., VectorNet)
- Neural map encoders
- Learned map embeddings

4. Modern Approaches to Motion Prediction

4.1 Transformer-Based Models

Transformers have become dominant due to their ability to model:

- Long temporal dependencies
- Multi-agent interactions
- Contextual reasoning

Examples include Motion Transformer (MTR) and Wayformer, which outperform traditional CNN and RNN approaches.

4.2 Multi-Modal Trajectory Prediction

Future trajectories are inherently uncertain. Modern methods predict multiple possible futures using:

- Probabilistic outputs
- Anchor-based trajectories
- Latent variable models

This improves robustness in real-world scenarios.

4.3 Diffusion Models for Trajectory Forecasting

Diffusion models represent a recent breakthrough by:

- Iteratively refining trajectory predictions
- Modeling complex uncertainty distributions
- Generating diverse realistic trajectories

They outperform GAN-based and deterministic methods in many benchmarks.

4.4 Interaction Modeling

Accurate prediction requires modeling interactions between agents. Techniques include:

- Graph Neural Networks
- Attention mechanisms
- Social pooling strategies

These allow models to capture dependencies such as yielding, overtaking, and collision avoidance.

5. Motion Planning and Closed-Loop Evaluation

Motion planning extends prediction by generating trajectories for the autonomous vehicle itself.

5.1 Open-Loop vs Closed-Loop

- **Open-loop:** Predict trajectories without affecting environment
- **Closed-loop:** Model interacts with environment dynamically

Closed-loop evaluation is more realistic but significantly more challenging.

5.2 Simulation-Based Benchmarks

Recent frameworks such as nuPlan provide:

- Realistic simulation environments
- Standardized evaluation metrics
- Scenario-based testing

These benchmarks reveal that performance often drops significantly in closed-loop settings, highlighting the importance of robust models.

6. Impact of Data Scale

Scaling dataset size continues to be a key factor in performance improvement. Key observations:

- Larger datasets improve generalization
- Rare scenarios become learnable
- Model performance scales with data size

Recent trends suggest moving toward:

- **Millions of driving hours**
- Synthetic data augmentation
- Simulation-generated scenarios

7. Discussion

Despite major progress, several challenges remain:

- Long-tail rare events (e.g., accidents)
- Domain adaptation across cities
- Real-time inference constraints
- Safety and interpretability

Future research directions include:

- Foundation models for autonomous driving
- End-to-end learning systems
- Integration of simulation and real-world data

8. Conclusion

In this paper, we revisited the role of large-scale datasets in motion prediction for autonomous driving and extended the discussion with recent advancements. Modern

approaches leveraging transformers, diffusion models, and large-scale datasets significantly improve prediction and planning performance.

However, achieving fully autonomous driving requires further scaling of both data and models, along with robust closed-loop evaluation. Continued efforts in dataset development, simulation, and learning paradigms will be essential for the next generation of autonomous systems.

References

- [1] A. Geiger et al., “KITTI Vision Benchmark Suite,” 2013.
- [2] M. Chang et al., “Argoverse Dataset,” CVPR, 2019.
- [3] Waymo, “Waymo Open Dataset,” 2019.
- [4] H. Caesar et al., “nuScenes Dataset,” 2020.
- [5] J. Houston et al., “One Thousand and One Hours,” CoRL, 2020.
- [6] Y. Chai et al., “MultiPath: Multiple Trajectory Prediction,” 2020.
- [7] Wayformer, “Motion Forecasting with Transformers,” 2022.
- [8] MTR: Motion Transformer, 2022.
- [9] nuPlan Benchmark, 2022.
- [10] Diffusion Models for Motion Prediction, 2023.