Uncertainty Modeling in Autonomous Vehicle Trajectory Prediction: A Comprehensive Survey

Anonymous CVPR submission

Paper ID *****

Abstract

Agent Behavior prediction is a critical component in autonomous driving systems, requiring the modeling of inherent uncertainties in an agent's future motion. This survey provides a comprehensive overview of uncertainty quantification approaches in agent behavior prediction, categorizing them into three main paradigms: probabilistic distribution-based models, generative models, and heatmap-based representations. We analyze how these paradigms address different aspects of uncertainty—including intent ambiguity, control variations, and inter-agent interactions—and evaluate their performance across standard benchmarks. Our comparison reveals the trade-offs between model expressiveness, computational efficiency, and deployment practicality. We conclude by identifying promising research directions that could advance uncertainty-aware trajectory prediction, ultimately contributing to safer and more reliable autonomous driving systems in complex real-world environments.

1. Introduction

Behavior prediction is a critical component to make self-driving cars work [26, 31]. In autonomous driving, behavior prediction forecasts the future trajectories and intentions of dynamic agents around the self-driving car. It is the process of forecasting how these dynamic agents, such as vehicles, pedestrians, or cyclists will move and interact based on historical observations and contextual cues from the environment. It serves as a critical link between perception (understanding the environment) and planning (determining the next actions of the self-driving car). Accurate behavior prediction is fundamental to autonomous vehicle safety and efficiency. Anticipating future agent actions allows self-driving cars to make proactive, risk-aware decisions, reducing accident likelihood and improving navigation in unpredictable urban environments

Future behavior prediction is inherently uncertain due to the stochastic nature of real-world environments and human decision-making. To illustrate this with an example, let us consider the uncertainty in future behavior prediction from the perspective of an agent navigating in the environment of an autonomous vehicle.

Intent Uncertainty This source of uncertainty comes from the fundamental ambiguity in the agent's intent. [5, 30, 38] For instance, if an agent is waiting at an intersection and originally planned to continue straight, they may suddenly decide to make a right turn because their final destination might change or because their intent changes. For example, an agent intending to travel to a grocery store may decide to make an additional stop at a gas station along the route, potentially changing their trajectory mid-journey and altering their final desired goal position. Modeling this uncertainty is inherently stochastic, meaning that even with perfect models and unlimited data, we can never achieve 100% prediction accuracy due to the random nature of human decision-making.

Control Uncertainty This source of uncertainty can arise even when the agent's intent is clear [5, 38]. Even if we are entirely certain that an agent is going to proceed straight, there remains growing uncertainty about its exact future pose. This uncertainty increases the further into the future we try to predict. Factors such as road conditions or subtle variations in pedal force applied by human drivers contribute to this growing pose uncertainty over time.

Interaction Uncertainty: This source of uncertainty arises from the complex interactions between multiple agents in a shared environment. For example, a vehicle that plans to accelerate may suddenly brake if another vehicle or pedestrian unexpectedly enters its path. Since agents operate within shared environments, their behaviors depend on responses from other agents, creating cascading uncertainties that make precise prediction of future behavior increasingly difficult as the number of interacting agents grows

From a machine-learning modeling perspective, these uncertainties can be categorized as *aleatoric* or *epistemic* [7, 10, 21, 22].

- **Aleatoric uncertainty** arises from the inherent stochasticity of the environment and represents randomness that exists regardless of data quantity or model sophistication. This type of uncertainty captures the unpredictable nature of real-world systems where identical initial conditions can lead to different outcomes. In autonomous driving, it manifests as the inherent unpredictability in human behavior that cannot be eliminated even with perfect modeling.

- **Epistemic uncertainty** stems from the model's lack of knowledge and can be reduced through improved architectures or increased data diversity. This uncertainty reflects limitations in model capacity or distributional shifts between training and deployment environments. It becomes particularly pronounced when encountering out-of-distribution samples, edge cases, or long-tail phenomena that were underrepresented in the training data.

Why is modeling uncertainty in behavior prediction critical for autonomous vehicle safety? Modeling future agent behaviors with uncertainty is a critical advancement in autonomous vehicle systems.

Earlier prediction systems produced only a single 'best' trajectory, often resulting in unrealistic interpolations that misled autonomous vehicle planners. For instance, when an agent had multiple valid choices (e.g., going straight or turning right), models sometimes generated an implausible diagonal path [1, 3, 16, 23, 28, 32]. This could lead to poor planning decisions, increasing the risk of unexpected lane deviations and potential collisions. This could result in a trajectory cutting across lanes or medians—one that no reasonable driver would actually follow. This limitation indicates that earlier models lacked the fundamental capability to represent behavioral uncertainty. The consequential impact emerged in how downstream systems in autonomous vehicles processed these predictions. When presented with only a single trajectory devoid of uncertainty information, the self-driving car could only perform binary collision detection against this ”average” path. The autonomous vehicle would then use this limited binary collision assessment to attempt planning a collision-free path, despite being based on potentially unrealistic agent predictions. This methodology introduces significant risk and represents a rudimentary approach to collision assessment. To compensate for these limitations, practical implementations required numerous heuristics to enhance robustness and accuracy of collision detection algorithms. The introduction of uncertainty modeling eliminates the need for these heuristic approaches. Probabilistic trajectory representations allow systems to shift from binary collision detection to a more informative 'probability of collision' assessment. This probabilistic framework provides autonomous vehicles with substantially improved capabilities for quantifying collision risks when interacting with multiple agents.

The growing importance of uncertainty-aware trajectory prediction in autonomous driving has led to significant research advances in recent years. Various approaches have emerged to address the fundamental challenge of representing and quantifying uncertainty in agent behavior. To systematically explore these developments, we present a comprehensive analysis of uncertainty modeling techniques in trajectory prediction. This work is structured around these key components:

- Section 2 presents an overview of different paradigms for quantifying uncertainty in trajectory prediction, including probabilistic distributions, generative approaches, and heatmap-based representations.
- Section 3 delivers an in-depth survey of various models in recent literature categorized by these uncertainty quantification approaches, analyzing their theoretical foundations, implementation techniques, and how they address both aleatoric and epistemic uncertainty in autonomous driving contexts.
- Sections 4 and 5 provide a quantitative comparison across benchmark datasets and outline promising future research directions, respectively.

Key Contributions:

- A structured taxonomy of uncertainty quantification paradigms for trajectory prediction.
- A comparative analysis of state-of-the-art models based on accuracy, uncertainty calibration, and deployment feasibility.
- Insights into open challenges and future research directions, focusing on hybrid modeling strategies and real-world integration.

2. Background on Uncertainty Quantification Paradigms

This section provides an overview of the major paradigms for quantifying uncertainty in behavior prediction.

2.1. Probabilistic Distribution-Based Models

Probabilistic models for behavior prediction typically represent an agent's intent uncertainty by defining a fixed set of possible futures the agent might choose. Each future corresponds to a distinct high-level decision (e.g., going straight, turning left), and a probability distribution is assigned over these modes.

A common approach is to predefine a set of trajectory anchors that serve as representative modes of future motion. Figure 1 demonstrates an example of this from the MultiPath [5] paper. Given an observed history, the model predicts the probability of each anchor being chosen, along with deviations from the anchor trajectory. This formulation provides a structured way to capture multi-modal uncertainty while maintaining a compact representation. [5, 29, 35]

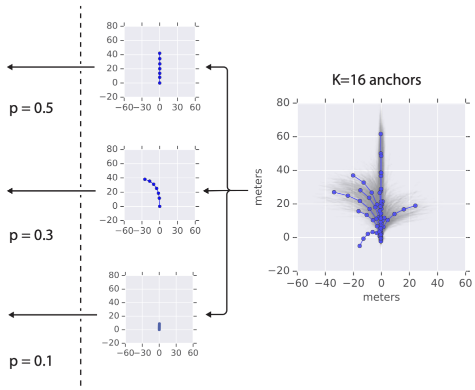

Figure 1. Probabilistic distribution-based models represent uncertainty through multiple trajectory anchors with associated probabilities. Left: Individual anchors with varying probability values (p). Right: Complete set of $K = 16$ anchors showing the multi-modal distribution. Figure adapted from [5].

Beyond intent uncertainty, some probabilistic models also explicitly model *control uncertainty*, which accounts for variations in trajectory execution even when the intent is fixed. This is typically modeled as a probability distribution over trajectory refinements, allowing the model to express deviations due to environmental factors, agent dynamics, or perception noise. [5]

Most probabilistic models follow a *marginal prediction* approach, where each agent's trajectory is predicted independently without explicitly modeling interactions. While this simplifies computation and training, it does not inherently capture interaction uncertainty, where the future motion of one agent is conditioned on that of others. Some probabilistic models attempt to address this limitation by incorporating interaction-aware mechanisms, but handling joint uncertainty remains an open challenge [27].

2.2. Generative Models for Uncertainty Modeling

Generative models for behavior prediction produce a distribution over possible future trajectories of an agent through sampling-based approaches. Typically, these models sample multiple seeds from a normal distribution to generate plausible futures that reflect both intent and control uncertainty. Since these models generate sample trajectories of all agents in the environment at once from a given seed, they provide interaction-aware predictions. This contrasts with most Probabilistic Distribution-Based Models, which generally only perform marginal trajectory prediction. Figure 2 demonstrates an example of this approach from the *Trajectron++* [33] paper which models multi-agent interactions between pedestrians, showing how multiple possible trajectories are predicted while accounting for inter-agent influences. [4, 14, 15, 20, 33, 34, 36]

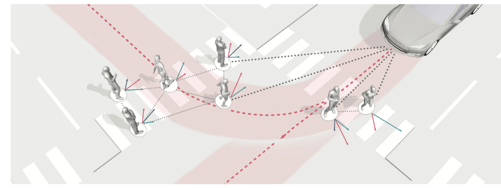

Figure 2. Generative models use structured representations to model interaction uncertainty, ensuring sampled trajectories remain consistent. Figure adapted from [33].

2.3. Heatmap-Based Representations

Heatmap-based models quantify uncertainty by predicting spatial probability distributions over future agent positions rather than discrete trajectory sets. Unlike probabilistic models that define explicit probability distributions or generative models that sample from a latent space, heatmap-based approaches produce a continuous probability field over a spatial grid. Each pixel in the heatmap represents a potential future position, with intensity values indicating the likelihood of occupancy at that location. This formulation naturally captures multi-modal uncertainty, as the spread of high-probability regions corresponds to the range of possible future motions. This is then followed by sampling from the heatmap to generate trajectories. Figure 3 demonstrates an example of this approach from the HOME paper which generates trajectory predictions by first creating spatial probability heatmaps and then sampling representative paths with associated confidence scores while accounting for interactions with other agents. [11–13, 19, 24]

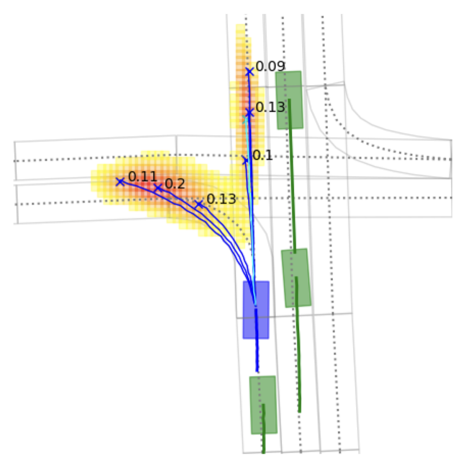

Figure 3. HOME model visualization showing probability heatmaps, sampled trajectories with confidence scores, and interacting agents at an intersection. Figure adapted from [11].

3. Taxonomy of Uncertainty Quantification Approaches in Behavior Prediction

This section provides an in-depth survey of various models in recent literature categorized by their uncertainty quantification paradigms. The focus is on analyzing their theoretical foundations, implementation techniques, and how they address both aleatoric and epistemic uncertainty in autonomous driving contexts.

3.1. Probabilistic Distribution-Based Models

MultiPath [5] models uncertainty in behavior prediction by representing future motion as a discrete distribution over predefined anchor trajectories with Gaussian-distributed offsets. It explicitly separates intent uncertainty from control uncertainty. Intent uncertainty is modeled as a discrete probability distribution over trajectory anchors. Control uncertainty is captured using normally distributed offsets at each timestep, where the mean represents deviation from the predefined anchor trajectory and the covariance represents control uncertainty.

MultiPath++ [35] extends MultiPath by refining how multimodal uncertainty is captured and aggregated. Instead of relying on static trajectory anchors, it learns latent anchor embeddings, improving adaptability to diverse driving contexts. Uncertainty is modeled using a Gaussian Mixture Model (GMM), where multiple predicted trajectories are assigned probabilities to represent stochastic human behavior. To further enhance uncertainty modeling, MultiPath++ employs an Expectation-Maximization (EM)-based clustering algorithm, which aggregates multiple predictor outputs into a structured representation of future motion. This clustering mechanism ensures that the final trajectory predictions remain diverse while maintaining high-confidence probability estimates, leading to more calibrated and reliable multimodal predictions compared to direct sampling-based approaches. This approach reduces one of the critical problems of intent mode collapse faced by MultiPath.

Target-driven trajectory prediction (TNT) [38] models uncertainty by explicitly structuring trajectory generation into two stages: target selection and motion estimation. In the first stage, the model predicts a probability distribution over discrete target locations, capturing intent uncertainty by representing the likelihood of different high-level decisions. Given a selected target, the second stage estimates continuous trajectories conditioned on that target, assuming a unimodal distribution per target. This decomposition enables clear separation between uncertainty in decision-making and uncertainty in execution, improving interpretability. The model further refines uncertainty representation by applying a ranking mechanism and non-maximum suppression-like filtering to eliminate redundant trajectory hypotheses while ensuring diverse, high-likelihood predictions.

Scene Transformer [27] explicitly models control uncertainty in multi-agent trajectory prediction by parameterizing each predicted trajectory with a Laplace distribution, allowing it to capture variations in motion at each timestep. Unlike models that predict independent agent trajectories, Scene Transformer ensures joint consistency by applying self-attention across agents and time steps, enabling more structured uncertainty propagation. To improve efficiency, it factorizes attention into temporal and agent-specific components, ensuring that uncertainty estimates remain coherent across interacting agents.

3.2. Generative Models

Social GAN [15] models uncertainty by generating multiple plausible future trajectories through adversarial learning. The generator produces trajectory samples, while the discriminator evaluates their realism based on social interactions. Social GAN uses a variety loss to encourage diverse predictions and prevent mode collapse. The diversity of sampled trajectories implicitly represents stochastic motion uncertainty, allowing the model to account for the variability in agent behavior while maintaining social compliance.

Trajectron++ [33] captures uncertainty using a Conditional Variational Autoencoder (CVAE), where a discrete latent variable encodes multiple plausible futures. By sampling from this latent distribution, the model captures both high-level intent uncertainty (decision-making variability) and low-level execution uncertainty (trajectory deviations). Instead of predicting deterministic future paths, Trajectron++ outputs a distribution over possible trajectories, where the spread of sampled trajectories reflects the model's confidence in its predictions. Additionally, uncertainty is propagated through probabilistic control distributions over acceleration and steering, ensuring dynamically feasible outputs.

FloMo (Flow-based Motion Prediction) [34] uses a normalizing flow-based approach to model uncertainty by treating motion prediction as a density estimation problem, learning a multimodal probability distribution over future trajectories. Instead of generating discrete trajectory sets, FloMo learns a direct transformation from a base noise distribution to the trajectory space, ensuring both diverse and probabilistically meaningful predictions. This approach enables efficient sampling of future trajectories while maintaining tractability in likelihood computation, making it particularly effective for uncertainty-aware motion planning. Additionally, FloMo introduces a likelihood-based training method, addressing common issues such as mode collapse in GANs and indirect uncertainty estimation in VAEs. To ensure stability, FloMo incorporates a noise-injection technique, which prevents likelihood spikes and enables robust generalization to real-world driving scenarios.

With the success of *Denoising Diffusion Probabilistic Models (DDPMs)* [18] in various domains, this approach has been extended into behavior prediction through *Motion-Diffuser* [20]. MotionDiffuser captures uncertainty through a diffusion-based generative process, where noisy initial trajectory samples are iteratively refined into structured predictions. Unlike GANs or VAEs, which generate trajectories in a single step, MotionDiffuser starts from a Gaussian noise distribution and progressively denoises it over multiple iterations, ensuring controlled and diverse trajectory generation. By conditioning this denoising process on past motion and scene context, MotionDiffuser produces multimodal trajectory distributions that reflect inherent uncertainty while maintaining physical plausibility. This iterative refinement enables fine-grained uncertainty representation.

3.3. Heatmap-Based Representations

Heatmap output for future motion estimation (HOME) [11] represents uncertainty in trajectory forecasting using a 2D probability heatmap, where each pixel encodes the likelihood of an agent occupying that position in the future. Unlike methods constrained to predefined trajectory clusters, HOME captures the full distribution of possible outcomes, supporting flexible multimodal predictions. To refine uncertainty representation, HOME integrates attention mechanisms that model inter-agent interactions, ensuring that predicted probability distributions reflect realistic behavioral variations. Additionally, it introduces two sampling strategies for trajectory extraction: one that prioritizes coverage to minimize the probability of missing the true future position and another that optimizes for final displacement accuracy. This enables a trade-off between trajectory diversity and precision without requiring retraining, making HOME well-suited for adaptive motion forecasting in autonomous driving applications.

Graph-oriented heatmap output for future motion estimation (GOHOME) [13] improves upon standard heatmap-based representations by incorporating lane connectivity information, ensuring that trajectory predictions align with real-world road constraints. Instead of computing a probability distribution over a free-space grid, GOHOME generates lane-level probability rasters and projects them onto a global heatmap, preserving road structure and traffic flow patterns. By ranking the most probable lanelets and computing heatmaps only for those regions, GOHOME maintains multimodal uncertainty representation while reducing computational overhead. Additionally, the model employs uncertainty-aware ensembling, combining multiple probability maps to enhance prediction accuracy while preventing mode collapse, a common issue in trajectory-based forecasting. This structured approach allows GOHOME to capture both aleatoric and epistemic uncertainty while maintaining computational efficiency, making it particularly effective for structured, long-horizon trajectory prediction.

Trajectory heatmap output with learned multi-agent sampling (THOMAS) [12] refines heatmap-based uncertainty modeling by using a hierarchical decoding process that iteratively refines predictions at multiple resolutions. Instead of generating a single high-resolution probability map, THOMAS begins with a coarse-grained heatmap to approximate the overall uncertainty distribution and then progressively refines it by focusing on high-probability regions. This hierarchical strategy improves both computational efficiency and prediction accuracy, concentrating resources where they are most needed. Additionally, THOMAS employs a deterministic sampling mechanism to extract diverse, multimodal trajectory endpoints from the heatmap, ensuring that different future intentions and motion variations are captured. To maintain scene consistency across interacting agents, THOMAS introduces a modality recombination module, which aligns agent trajectories to prevent collisions and ensure that predictions remain jointly coherent. This structured uncertainty modeling approach makes THOMAS highly effective for multi-agent trajectory prediction, where maintaining scene-level consistency is crucial for realistic and safe autonomous behavior.

4. Model Comparison and Analysis

This section presents a qualitative and quantitative comparison of uncertainty-aware trajectory prediction models across standard benchmark datasets, evaluating their performance using standardized metrics that assess both accuracy and uncertainty calibration.

4.1. Qualitative Analysis

We propose a comprehensive framework for evaluating trajectory prediction paradigms across five critical dimensions that capture both the quality of uncertainty representation and practical system considerations:

- **Intent Uncertainty**: How well the approach captures the fundamental ambiguity in agent goals and intentions
- **Control Uncertainty**: The model's capacity to represent variability in trajectory execution given a fixed intent
- **Interaction Uncertainty**: How effectively the approach captures the interdependence between multiple agents' trajectories
- **Computational Efficiency**: The relative computational resources required for inference
- **Deployment Practicality**: Suitability for real-world implementation in autonomous systems

Table 1 presents our analysis of how the three major paradigms—probabilistic models, generative approaches, and heatmap-based methods—perform across these dimensions. Probabilistic models provide structured uncertainty estimation but may struggle with modeling complex interactions. Generative approaches offer flexible and interaction-aware predictions but suffer from stochasticity and mode coverage issues, including intent mode collapse (where the model fails to capture distinct, meaningful intent modes). Heatmap-based methods provide highly multimodal spatial uncertainty representation but introduce computational complexity and require additional trajectory extraction steps. The choice of approach depends on system requirements, and future work may focus on hybrid strategies that combine the strengths of multiple paradigms.

4.2. Quantitative Analysis

The trajectory prediction field utilizes several benchmark datasets for evaluation, with Argoverse [6] being one of the most widely used due to its rich map data and diverse urban scenarios. Other notable datasets include Waymo Open Motion Dataset [8], nuScenes [2], and INTERACTION [37], each offering unique characteristics for specialized testing scenarios. For our analysis, we focus on Argoverse, which serves as a tracking benchmark with over 30K scenarios from Pittsburgh and Miami, sampled at 10 Hz, where each sequence contains an agent whose future locations must be predicted over a 3-second horizon.

To systematically evaluate model performance on this benchmark, we consider both accuracy and uncertainty quality metrics. Accuracy metrics measure how closely predicted trajectories align with actual movements:

- **Minimum Final Displacement Error (minFDE)** quantifies the final position error by selecting the closest prediction among multiple outputs.
- **Minimum Average Displacement Error (minADE)** computes the average deviation throughout the trajectory, capturing overall path alignment.
- **Miss Rate (MR)** highlights critical prediction failures by counting instances where the predicted trajectory deviates beyond a predefined threshold, typically 2 meters for vehicles.

Uncertainty quality metrics assess how well models express confidence in their predictions:

- **Brier-minFDE** integrates accuracy and calibration, rewarding models whose confidence matches actual performance.
- **Diversity-Aware Coverage (DAC)** evaluates how comprehensively predictions capture the range of plausible future motions, encouraging models to account for multiple possible outcomes rather than only the most probable one.

Table 2 presents the performance of various uncertainty-aware trajectory prediction models on the Argoverse Benchmark test dataset. HOME+GOHOME achieves the lowest miss rate (0.08), demonstrating superior performance in capturing the most likely trajectories. MultiPath++ shows the best overall accuracy with the lowest minFDE (1.21) and minADE (0.79), suggesting strong general prediction capability. Most models achieve near-perfect diversity coverage (DAC = 0.99) except for Social CVAE, which shows significantly worse performance across all metrics. This may indicate fundamental limitations in the pure CVAE approach when applied to complex driving scenarios.

It's worth noting that performance differences may also relate to model complexity and computational resources, not just algorithmic superiority. Models with larger parameter counts or more sophisticated architectures may achieve better results due to increased model capacity rather than fundamental advancements in uncertainty modeling. Our analysis focuses primarily on the uncertainty representation capabilities rather than architectural details.

5. Conclusion and Future Directions

This survey has provided a comprehensive overview of uncertainty modeling approaches in trajectory prediction for autonomous driving. We have categorized these approaches into three main paradigms—probabilistic distribution-based models, generative models, and heatmap-based representations—and analyzed their theoretical foundations and implementation techniques. Each paradigm has strengths, but deployment challenges persist, particularly in ensuring real-time inference under hardware constraints, handling edge cases, and maintaining robustness in dynamic traffic environments. Addressing these issues is crucial for translating uncertainty-aware trajectory prediction from research to production-level autonomous systems. Probabilistic models provide structured representations but often struggle with complex interactions; generative approaches excel at capturing rich interaction dynamics but face deployment challenges; and heatmap-based methods offer intuitive spatial uncertainty representations but require additional processing steps. [9, 17, 25]

Future research should develop hybrid approaches that integrate the interpretability of probabilistic models, the flexibility of generative methods, and the multimodal richness of heatmap-based approaches. Additionally, new benchmarks should evaluate models not only on accuracy and uncertainty calibration but also on real-time efficiency and adaptability to domain shifts. Addressing these aspects is essential for translating uncertainty-aware trajectory prediction into real-world autonomous systems. Additionally, the field would benefit from standardized evaluation protocols specifically designed to assess uncertainty quality beyond traditional accuracy metrics. Addressing these challenges will require interdisciplinary efforts spanning machine learning, robotics, and human behavior modeling, ultimately enabling autonomous vehicles to safely navigate the complexity and unpredictability of real-world environments. As uncertainty modeling techniques continue to evolve, their integration into end-to-end autonomous

Paradigm	Intent Uncertainty	Control Uncertainty	Interaction Uncertainty	Computational Efficiency	Deployment Practicality
Probabilistic Models	**High** Explicitly modeled using discrete trajectory modes	**Moderate** Gaussian offsets capture variability, lacks fine-grained execution modeling	**Low** Typically models agents independently, leading to interaction inconsistency	**High** Efficient due to discrete mode selection, low inference cost	**High** Widely used in real-world AV systems
Generative Models	**Moderate** Latent variables encode diverse intents, mode coverage can be inconsistent	**High** Well-suited for stochastic motion modeling	**High** Jointly models agent interactions, producing socially consistent trajectories	**Low** Computationally expensive due to iterative sampling	**Low** Challenges for real-time deployment due to latency
Heatmap-Based Models	**High** Predicts spatial probability distributions, naturally captures intent uncertainty	**Moderate** Some methods refine uncertainty via probability smoothing	**High** Well-suited for multi-agent interactions through shared spatial grids	**Moderate** Memory-intensive, requires trajectory post-processing	**Moderate** Used in research, but deployment remains challenging

Table 1. Comparison of uncertainty quantification paradigms in behavior prediction.

Model	minFDE	minADE	MR	Brier-minFDE	DAC
MultiPath++	1.21	0.79	0.13	1.79	0.99
TNT	1.45	0.91	0.17	2.14	0.99
Scene Transformer	1.23	0.80	0.13	1.89	0.99
HOME + GOHOME	1.29	0.89	**0.08**	1.86	0.98
THOMAS	1.44	0.94	0.10	1.97	0.98
Social CVAE	3.23	2.49	0.39	4.27	0.67

Table 2. Performance comparison of uncertainty-aware trajectory prediction models on Argoverse [6].

driving systems will play a crucial role in advancing the safety, reliability, and performance of self-driving technologies.

References

[1] Mayank Bansal, Alex Krizhevsky, and Abhijit Ogale. Chauffeurnet: Learning to drive by imitating the best and synthesizing the worst. *arXiv preprint arXiv:1812.03079*, 2018. 2

[2] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11621–11631, 2020. 6

[3] Sergio Casas, Wenjie Luo, and Raquel Urtasun. Intentnet: Learning to predict intention from raw sensor data. In *Conference on Robot Learning*, pages 947–956. PMLR, 2018. 2

[4] Sergio Casas, Cole Gulino, Simon Suo, and Raquel Urtasun. The importance of prior knowledge in precise multimodal prediction. In *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pages 2295–2302. IEEE, 2020. 3

[5] Yuning Chai, Benjamin Sapp, Mayank Bansal, and Dragomir Anguelov. Multipath: Multiple probabilistic anchor trajectory hypotheses for behavior prediction. *arXiv preprint arXiv:1910.05449*, 2019. 1, 2, 3, 4

[6] Ming-Fang Chang, John Lambert, Patsorn Sangkloy, Jagjeet Singh, Slawomir Bak, Andrew Hartnett, De Wang, Peter Carr, Simon Lucey, Deva Ramanan, et al. Argoverse: 3d tracking and forecasting with rich maps. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8748–8757, 2019. 6, 7

[7] Armen Der Kiureghian and Ove Ditlevsen. Aleatory or epistemic? does it matter? *Structural safety*, 31(2):105–112, 2009. 1

[8] Scott Ettinger, Shuyang Cheng, Benjamin Caine, Chenxi Liu, Hang Zhao, Sabeek Pradhan, Yuning Chai, Ben Sapp, Charles R Qi, Yin Zhou, et al. Large scale interactive motion forecasting for autonomous driving: The waymo open motion dataset. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9710–9719, 2021. 6

[9] Angelos Filos, Panagiotis Tigkas, Rowan McAllister, Nicholas Rhinehart, Sergey Levine, and Yarin Gal. Can autonomous vehicles identify, recover from, and adapt to distribution shifts? In *International Conference on Machine Learning*, pages 3145–3153. PMLR, 2020. 6

[10] Yarin Gal et al. Uncertainty in deep learning. 2016. 1

[11] Thomas Gilles, Stefano Sabatini, Dzmitry Tsishkou, Bogdan Stanciulescu, and Fabien Moutarde. Home: Heatmap output for future motion estimation. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 500–507. IEEE, 2021. 3, 5

[12] Thomas Gilles, Stefano Sabatini, Dzmitry Tsishkou, Bogdan Stanciulescu, and Fabien Moutarde. Thomas: Trajectory heatmap output with learned multi-agent sampling. *arXiv preprint arXiv:2110.06607*, 2021. 5

[13] Thomas Gilles, Stefano Sabatini, Dzmitry Tsishkou, Bogdan Stanciulescu, and Fabien Moutarde. Gohome: Graph-oriented heatmap output for future motion estimation. In *2022 international conference on robotics and automation (ICRA)*, pages 9107–9114. IEEE, 2022. 3, 5

[14] Junru Gu, Chen Sun, and Hang Zhao. Densetnt: End-to-end trajectory prediction from dense goal sets. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 15303–15312, 2021. 3

[15] Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. Social gan: Socially acceptable trajectories with generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2255–2264, 2018. 3, 4

[16] Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282, 1995. 2

[17] Dan Hendrycks and Kevin Gimpel. A baseline for detecting misclassified and out-of-distribution examples in neural networks. *arXiv preprint arXiv:1610.02136*, 2016. 6

[18] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020. 5

[19] Xin Huang, Stephen G McGill, Jonathan A DeCastro, Luke Fletcher, John J Leonard, Brian C Williams, and Guy Rosman. Diversitygan: Diversity-aware vehicle motion prediction via latent semantic sampling. *IEEE Robotics and Automation Letters*, 5(4):5089–5096, 2020. 3

[20] Chiyu Jiang, Andre Cornman, Cheolho Park, Benjamin Sapp, Yin Zhou, Dragomir Anguelov, et al. Motiondiffuser: Controllable multi-agent motion prediction using diffusion. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9644–9653, 2023. 3, 5

[21] Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in neural information processing systems*, 30, 2017. 1

[22] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, 30, 2017. 1

[23] Wenjie Luo, Bin Yang, and Raquel Urtasun. Fast and furious: Real time end-to-end 3d detection, tracking and motion forecasting with a single convolutional net. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 3569–3577, 2018. 2

[24] Karttikeya Mangalam, Harshayu Girase, Shreyas Agarwal, Kuan-Hui Lee, Ehsan Adeli, Jitendra Malik, and Adrien Gaidon. It is not the journey but the destination: Endpoint conditioned trajectory prediction. In *Computer vision–ECCV 2020: 16th European conference, Glasgow, UK, August 23–28, 2020, proceedings, part II 16*, pages 759–776. Springer, 2020. 3

[25] Rhiannon Michelmore, Matthew Wicker, Luca Laurenti, Luca Cardelli, Yarin Gal, and Marta Kwiatkowska. Uncertainty quantification with statistical guarantees in end-to-end autonomous driving control. In *2020 IEEE international conference on robotics and automation (ICRA)*, pages 7344–7350. IEEE, 2020. 6

[26] Sajjad Mozaffari, Omar Y Al-Jarrah, Mehrdad Dianati, Paul Jennings, and Alexandros Mouzakitis. Deep learning-based vehicle behavior prediction for autonomous driving applications: A review. *IEEE Transactions on Intelligent Transportation Systems*, 23(1):33–47, 2020. 1

[27] Jiquan Ngiam, Benjamin Caine, Vijay Vasudevan, Zhengdong Zhang, Hao-Tien Lewis Chiang, Jeffrey Ling, Rebecca Roelofs, Alex Bewley, Chenxi Liu, Ashish Venugopal, et al. Scene transformer: A unified architecture for predicting multiple agent trajectories. *arXiv preprint arXiv:2106.08417*, 2021. 3, 4

[28] Stefano Pellegrini, Andreas Ess, Konrad Schindler, and Luc Van Gool. You'll never walk alone: Modeling social behavior for multi-target tracking. In *2009 IEEE 12th international conference on computer vision*, pages 261–268. IEEE, 2009. 2

[29] Tung Phan-Minh, Elena Corina Grigore, Freddy A Boulton, Oscar Beijbom, and Eric M Wolff. Covernet: Multimodal behavior prediction using trajectory sets. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14074–14083, 2020. 2

[30] Nicholas Rhinehart, Rowan McAllister, Kris Kitani, and Sergey Levine. Precog: Prediction conditioned on goals in visual multi-agent settings. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 2821–2830, 2019. 1

[31] Andrey Rudenko, Luigi Palmieri, Michael Herman, Kris M Kitani, Dariu M Gavrila, and Kai O Arras. Human motion trajectory prediction: A survey. *The International Journal of Robotics Research*, 39(8):895–935, 2020. 1

[32] Amir Sadeghian, Ferdinand Legros, Maxime Voisin, Ricky Vesel, Alexandre Alahi, and Silvio Savarese. Car-net: Clairvoyant attentive recurrent network. In *Proceedings of the European conference on computer vision (ECCV)*, pages 151–167, 2018. 2

[33] Tim Salzmann, Boris Ivanovic, Punarjay Chakravarty, and Marco Pavone. Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII 16*, pages 683–700. Springer, 2020. 3, 4

[34] Christoph Schöller and Alois Knoll. Flomo: Tractable motion prediction with normalizing flows. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7977–7984. IEEE, 2021. 3, 4

[35] Balakrishnan Varadarajan, Ahmed Hefny, Avikalp Srivastava, Khaled S Refaat, Nigamaa Nayakanti, Andre Cornman, Kan Chen, Bertrand Douillard, Chi Pang Lam, Dragomir Anguelov, et al. Multipath++: Efficient information fusion and trajectory aggregation for behavior prediction. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 7814–7821. IEEE, 2022. 2, 4

[36] Ye Yuan and Kris Kitani. Diverse trajectory forecasting with determinantal point processes. *arXiv preprint arXiv:1907.04967*, 2019. 3

[37] Wei Zhan, Liting Sun, Di Wang, Haojie Shi, Aubrey Clausse, Maximilian Naumann, Julius Kümmerle, Hendrik Königshof, Christoph Stiller, Arnaud de La Fortelle, and Masayoshi Tomizuka. INTERACTION Dataset: An INTERnational, Adversarial and Cooperative moTION Dataset in Interactive Driving Scenarios with Semantic Maps. *arXiv:1910.03088 [cs, eess]*, 2019. 6

[38] Hang Zhao, Jiyang Gao, Tian Lan, Chen Sun, Ben Sapp, Balakrishnan Varadarajan, Yue Shen, Yi Shen, Yuning Chai, Cordelia Schmid, et al. Tnt: Target-driven trajectory prediction. In *Conference on Robot Learning*, pages 895–904. PMLR, 2021. 1, 4