ISSN: 2519-9439

Interaction-Field Residual Learning for Calibrated Lane-Change Intent under Mixed Traffic

[1]Anele Mthembu, [2]Thabo Pretorius

[1]Department of Computer Science, Faculty of Science, University of Cape Town, Rondebosch, Western Cape, South Africa

[2]Department of Computer Science and Applied Mathematics, School of Computer Science and Applied Mathematics, University of the Witwatersrand, Johannesburg, Gauteng, South Africa

Abstract

The prediction of lane-change intent in mixed traffic environments, characterized by the coexistence of human-driven vehicles and autonomous vehicles, presents a formidable challenge due to the stochastic nature of human driving behaviors and complex vehicular interactions. Conventional trajectory prediction and intent recognition models often struggle to capture the nuanced, continuous spatial-temporal dynamics of such environments, frequently resulting in overconfident or poorly calibrated predictions. To address these critical shortcomings, this paper proposes a novel framework termed Interaction-Field Residual Learning for Calibrated Lane-Change Intent. The framework introduces a continuous interaction-field representation that models traffic participants as sources of physical potentials, thereby capturing both repulsive and attractive social forces in a unified topographical space. Built upon this representation, a residual learning architecture is deployed to separate the deterministic vehicle dynamics from the highly stochastic, socially driven residual movements. Furthermore, a post-hoc calibration module is integrated to ensure that the probabilistic outputs accurately reflect the true likelihood of lane-change maneuvers, reducing critical false-positive rates in safety-critical scenarios. Extensive evaluations on large-scale naturalistic driving datasets demonstrate that the proposed method significantly outperforms existing baseline models in both predictive accuracy and uncertainty calibration metrics. The findings offer profound implications for the design of safe, socially aware navigation systems in transitional mixed-traffic periods.

Keywords: Lane-change intent, mixed traffic, interaction field, residual learning, uncertainty calibration

1.Introduction

1.1 Background and Motivation

The transition toward fully autonomous transportation systems is anticipated to span several decades, giving rise to an extended transitional phase known as mixed traffic. In this environment, human-driven vehicles and autonomous vehicles must share the same road space, necessitating advanced perception and prediction capabilities from the autonomous agents [1]. The primary difficulty in navigating mixed traffic lies in the fundamental asymmetry of behaviors: human drivers exhibit a high degree of stochasticity, variability, and bounded rationality, whereas autonomous vehicles typically operate on deterministic, rule-based, or optimally derived control policies [2, 3]. This behavioral mismatch is particularly evident during complex maneuvers such as lane changing, which

ISSN: 2519-9439

require intense spatial-temporal coordination and implicit negotiation among multiple traffic participants. Accurate prediction of lane-change intent is therefore recognized as a cornerstone capability for ensuring the safety and efficiency of autonomous driving systems. When an autonomous vehicle approaches a slower-moving vehicle or navigates a congested highway, it must continuously infer the intentions of surrounding human drivers to plan its own trajectory proactively [4]. Failure to accurately predict a lane-change maneuver can lead to sudden, uncomfortable braking, disruption of traffic flow, or catastrophic collisions [5]. Despite considerable advancements in sensor technologies and machine learning algorithms, the accurate and timely prediction of lane-change intent remains an open problem. The core challenge stems from the fact that lane changing is not an isolated event but rather the culmination of prolonged, dynamic interactions between the ego vehicle and its surrounding environment [6, 7]. These interactions are heavily influenced by unobservable factors, including the driver's internal state, varying levels of aggression, and idiosyncratic responses to local traffic density.

1.2 Problem Statement

Existing approaches to lane-change intent prediction largely rely on discrete representations of the traffic state, often treating surrounding vehicles as independent entities represented by simple feature vectors such as position, velocity, and heading angle [8]. While these discrete representations are computationally efficient, they fail to capture the continuous, wave-like propagation of influence that vehicles exert on one another. For example, a sudden deceleration by a leading vehicle creates a ripple effect of cautious behavior that permeates the surrounding traffic, altering the probability of lane changes for vehicles further upstream [9, 10]. Methods that rely solely on pairwise distances or graph-based topologies often lose the granularity of these continuous spatial dynamics. Furthermore, many deep learning models applied to this domain, such as recurrent neural networks and standard convolutional architectures, output uncalibrated probabilities [11]. In the context of safety-critical applications, a model's confidence in its prediction is just as important as the prediction itself. If a model predicts a lane change with high confidence when the true probability is low, the autonomous vehicle might take unnecessary evasive actions, leading to phantom traffic jams or secondary accidents [12]. Conversely, under-confident predictions may delay necessary safety interventions. Thus, there is an urgent need for models that not only capture complex interaction dynamics but also provide highly calibrated uncertainty estimates.

1.3 Contributions

To overcome these limitations, this paper introduces a comprehensive framework named Interaction-Field Residual Learning. The primary contributions of this research are threefold. First, we formulate a continuous interaction-field representation of the mixed traffic environment, mapping discrete vehicle states into a continuous potential field that elegantly captures the attractive and repulsive forces governing human driving behavior [13, 14]. Second, we develop a novel residual learning architecture that explicitly separates the deterministic kinematic progression of a vehicle from the stochastic, interaction-driven residuals. This architectural choice significantly enhances the model's ability to learn complex social behaviors without being overwhelmed by standard vehicle dynamics [15]. Third, we incorporate a rigorous uncertainty calibration mechanism that adjusts the raw logits of the neural network, ensuring that the predicted probabilities faithfully reflect empirical maneuver likelihoods.

2. Literature Review

2.1 Lane-Change Intent Prediction Methods

The literature on lane-change intent prediction can be broadly categorized into rule-based, classical machine learning, and deep learning approaches. Rule-based models, such as the widely studied Intelligent Driver Model and the Minimizing Overall Braking Induced by Lane change model, rely on predefined thresholds for safety and incentive criteria [16]. While highly interpretable and computationally lightweight, these models often fail to generalize across diverse traffic conditions because their parameters are heavily dependent on specific localized calibrations. Classical machine learning techniques, including Support Vector Machines, Hidden Markov Models, and Random

ISSN: 2519-9439

Forests, offered the next evolutionary step by learning lane-change triggers directly from data [17, 18]. However, these methods typically require exhaustive hand-crafted feature engineering, struggling to automatically extract high-level spatio-temporal dependencies from raw trajectory data. In recent years, deep learning has dominated the field of intent prediction. Recurrent Neural Networks, particularly Long Short-Term Memory networks, have been extensively employed to capture the temporal evolution of vehicle trajectories leading up to a lane change [19]. Researchers have demonstrated that feeding sequences of vehicle states into these networks yields superior predictive accuracy compared to static classifiers. Despite their temporal modeling prowess, standard recurrent models treat the traffic scene as a collection of independent sequences, ignoring the critical spatial relationships between adjacent vehicles [20, 21]. To mitigate this, subsequent studies introduced spatial pooling mechanisms and convolutional layers, though often at the cost of assuming a rigid grid structure for the dynamic road environment.

2.2 Interaction Modeling in Mixed Traffic

Modeling interactions in mixed traffic is inherently more complex than homogeneous autonomous environments because human drivers do not follow strict communication protocols. Graph Convolutional Networks have recently gained traction as a powerful tool for modeling vehicular interactions [22]. By representing vehicles as nodes and their interactions as edges, these networks can dynamically aggregate information from a vehicle's local neighborhood. However, graph-based methods face challenges in defining the edge weights, which are typically based on simple Euclidean distances or pre-computed heuristics that do not reflect the true, asymmetric nature of driving influence [23]. For instance, a vehicle directly in front exerts a significantly stronger influence on a driver than a vehicle situated at the same distance behind them. An alternative paradigm involves the use of social pooling or attention mechanisms, which learn to assign varying degrees of importance to different surrounding vehicles dynamically [24, 25]. While attention mechanisms have proven effective, they operate on discrete state vectors. Our proposed interaction field diverges from these discrete methods by treating the traffic environment as a continuous scalar and vector field. Drawing inspiration from potential field theory used in classical robotics, we construct a dense representation where the spatial influence of each vehicle decays exponentially with distance and relative velocity, creating a topographic map of driving risk and opportunity [26]. This continuous field naturally handles the asymmetric and non-linear interactions characteristic of mixed traffic without requiring explicit graph construction.

2.3 Residual Learning and Uncertainty Calibration

Residual learning, originally popularized in computer vision to facilitate the training of extremely deep networks, operates on the principle of learning the difference between the desired mapping and the identity mapping. In the context of trajectory and intent prediction, residual learning can be adapted to separate base kinematics from interaction-induced deviations [27, 28]. Some recent studies have utilized residual blocks within their temporal convolutional networks to improve gradient flow, but few have explicitly aligned the residual function with the stochastic components of driving behavior [29]. Our framework explicitly partitions the predictive task, utilizing a physics-based model for the base trajectory and a neural network solely for the interaction-driven residual intent.

Furthermore, the calibration of neural network outputs is a notoriously neglected area in autonomous driving research. Modern deep neural networks tend to be overconfident, producing probability distributions that are highly skewed [30, 31]. Techniques such as temperature scaling, Platt scaling, and evidential deep learning have been proposed to align predicted probabilities with actual outcomes in classification tasks. In safety-critical domains like lane-change prediction, ensuring that a predicted ninety percent probability of a lane change genuinely translates to a maneuver occurring ninety percent of the time is paramount [32]. This paper integrates an advanced temperature scaling mechanism specifically optimized for the highly imbalanced nature of lane-change events, where straight-driving significantly outnumbers lane-changing maneuvers.

ISSN: 2519-9439

3. Problem Formulation and Traffic Interaction Field

3.1 Traffic State Representation

The problem of lane-change intent prediction requires a robust mathematical formulation of the dynamic traffic environment. Let the traffic scene at time step t be defined by a set of vehicles, where one vehicle is designated as the target or ego vehicle, and the others constitute the surrounding interactive context. Each vehicle is characterized by a state vector encompassing its two-dimensional position coordinates, longitudinal and lateral velocities, acceleration components, and heading angle relative to the lane center [33]. In a mixed traffic environment, these state vectors exhibit high variance due to the varying control strategies employed by human drivers and autonomous agents. The historical trajectory of the target vehicle over a predefined observation window is utilized as the primary input for predicting the discrete lane-change intention at a future time horizon. The intention is typically modeled as a categorical variable, denoting left lane change, right lane change, or lane keeping.

Table 1Summary of Mathematical Notations

Symbol	Definition	Data Type	Range/Domain
t	Current time step	Integer	Greater than zero
H	Historical observation window	Integer	Arbitrary temporal span
p	Position vector of a vehicle	Continuous	Cartesian coordinate space
v	Velocity vector of a vehicle	Continuous	Real numbers
E	Interaction field potential	Continuous	Non-negative real numbers
y	Lane-change intent label	Categorical	Left, Right, Keep

To accurately process this information, the raw trajectory data must be transformed into a framework that inherently captures spatial dependencies. Traditional methods concatenate the state vectors of the ego vehicle and its nearest neighbors into a single, high-dimensional feature vector. However, this simplistic concatenation imposes an artificial ordering on the surrounding vehicles, which can confuse the learning algorithm when vehicles enter or leave the immediate vicinity [34, 35]. Instead, our approach transforms the discrete vehicular states into a unified spatial representation before applying sequence modeling techniques, ensuring permutation invariance and spatial continuity.

3.2 Interaction-Field Modeling

The core innovation of the proposed methodology is the construction of a dynamic Interaction Field. We hypothesize that every vehicle in a mixed traffic environment emanates a virtual field of influence that affects the decision-making processes of neighboring drivers. This field is analogous to an electromagnetic or gravitational field but is specifically tailored to the asymmetric realities of human driving. The field is composed of two primary components: a repulsive potential that discourages vehicles from occupying the same physical space and an attractive potential that encourages vehicles to align their speeds and maintain optimal traffic flow. The continuous interaction field potential at any spatial point is computed by aggregating the influences of all surrounding vehicles. The influence of a specific vehicle on a given point in space decays non-linearly based on the Euclidean distance and the relative velocity. For instance, a vehicle moving at a high velocity generates a forward-elongated field of repulsion, representing the extended braking distance required in front of it. Conversely, the lateral field decays more sharply, reflecting the constraints of lane boundaries. This interaction field provides a dense, continuous heatmap of the driving environment, naturally encoding the complex social forces at play without requiring explicit pair-wise calculations.

$$E_i(t) = \sum_{j \in N_i} \frac{\alpha}{Vertp_i - p_j Vert^2} + \beta exp(-\gamma Vert v_i - v_j Vert)$$

ISSN: 2519-9439

The formula above illustrates the fundamental calculation of the interaction field potential, where alpha, beta, and gamma represent learnable scaling parameters that dictate the strength and decay rate of the spatial and kinetic influences. By applying this formulation across a discretized grid surrounding the target vehicle, we generate a sequence of interaction field tensors. These tensors serve as the primary spatial input to our deep learning architecture, providing a rich, contextualized view of the environment that captures both the immediate threats and the broader traffic flow dynamics.

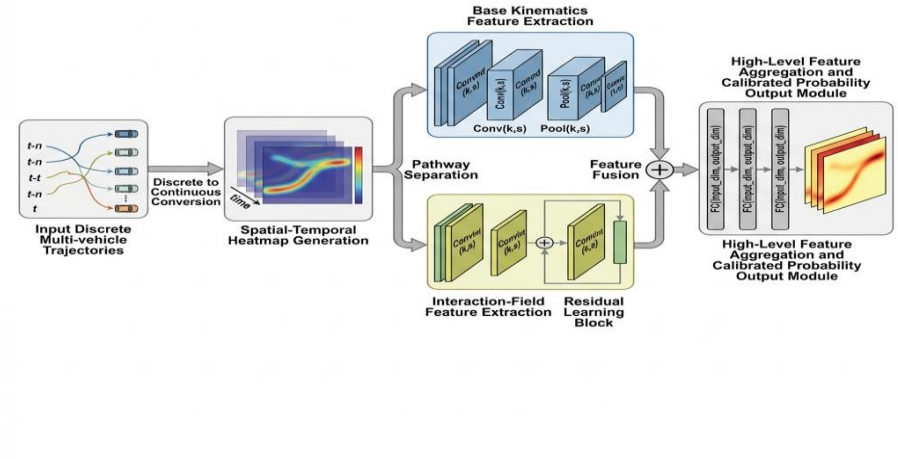


Figure 1Comprehensive Schematic of Interaction

4. Methodology: Residual Learning Framework

4.1 Core Architecture Design

The architecture of the Interaction-Field Residual Learning framework is designed to exploit the physical properties of the traffic environment while leveraging the representational capacity of deep neural networks. The system is composed of two parallel streams. The first stream is the deterministic kinematic module, which projects the future state of the vehicle based strictly on its current physical parameters using constant turn rate and velocity models. This stream provides a solid baseline prediction, assuming that the vehicle operates in a vacuum without any external influences. The second stream is the interaction-residual module, which processes the complex sequence of interaction field tensors to predict the deviations from the deterministic baseline caused by social traffic forces. The interaction-residual module employs a combination of Convolutional Neural Networks to extract spatial features from the interaction field tensors and Gated Recurrent Units to model the temporal evolution of these features over the observation window. By separating the prediction task in this manner, the neural network is relieved of the burden of learning basic physics. Instead, it can focus entirely its parameters on understanding the subtle, non-linear human behaviors that precede a lane change. For instance, the slight lateral nudging a driver performs when searching for a gap in the adjacent lane is captured purely as a residual signal against the straight-driving kinematic baseline.

$$haty_{t+1} = F(x_t; \theta) + R(h_t; \varphi) + \varepsilon$$

In the formulated update rule above, the final predicted intent is a composition of the base kinematic function parameterized by theta and the residual interaction function parameterized by phi, with an additional noise term epsilon representing irreducible environmental stochasticity. The outputs of both streams are eventually fused through a fully connected layer that maps the combined features into raw intent logits. This architectural separation ensures that even in scenarios where the interaction field is sparsely populated, the model falls back gracefully on robust kinematic projections, preventing erratic predictions common in end-to-end black-box models.

ISSN: 2519-9439

Table 2Comparative Analysis of State-of-the-Art Architecture Paradigms

Architecture Paradigm	Spatial Modeling Approach	Temporal Processing	Calibration Strategy
Standard Recurrent	Concatenated Feature Vectors	Long Short-Term Memory	None
Graph Convolutional	Node-Edge Heuristic Graphs	Temporal Graph Networks	Static Thresholding
Attention-Based	Discrete Multi-Head Attention	Transformer Encoder	Dropout Approximation
Proposed Interaction-Field	Continuous Potential Heatmaps	Gated Recurrent Units	Dynamic Temperature Scaling

4.2 Calibrated Uncertainty Estimation

While accurately predicting the intent is crucial, quantifying the confidence of that prediction is equally vital for safe autonomous navigation. To address the pervasive issue of neural network overconfidence, we integrate a post-hoc calibration module utilizing an advanced form of temperature scaling. Standard temperature scaling involves dividing the logits by a single, global scalar before applying the softmax function. However, because mixed traffic lane changes are highly contextual and classes are severely imbalanced, a global scalar is insufficient. Our framework introduces a dynamic, context-aware temperature scaling mechanism where the scaling factor is not a static parameter but is dynamically regressed based on the density and volatility of the current interaction field. In highly complex, densely packed traffic scenarios where multiple vehicles are interacting simultaneously, the model automatically increases the temperature parameter. This action softens the probability distribution, reflecting higher epistemic uncertainty. Conversely, in sparse traffic conditions where a vehicle's intent to change lanes is unambiguous, the temperature parameter is lowered, producing sharper, more decisive probabilities. This dynamic calibration ensures that the output probabilities are highly correlated with the empirical accuracy across various traffic complexities. The calibration process is optimized using a negative log-likelihood objective on a held-out validation set, ensuring that the model's predictive accuracy is not compromised during the calibration phase.

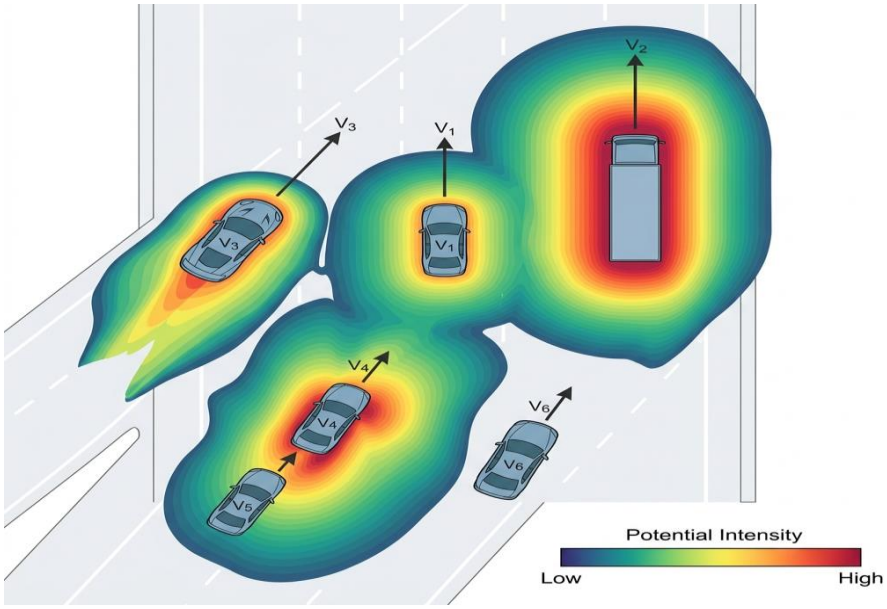


Figure 2: Dynamic Interaction Field Heatmap Visualization

ISSN: 2519-9439

5. Experimental Setup and Data Acquisition

5.1 Dataset Description

To rigorously evaluate the proposed Interaction-Field Residual Learning framework, extensive experiments were conducted using two prominent large-scale naturalistic driving datasets. These datasets were specifically selected for their high-fidelity trajectory recordings and their comprehensive coverage of complex, multi-vehicle interactions on highways and urban expressways. The data encompasses thousands of hours of driving behavior, capturing a wide array of traffic densities, varying weather conditions, and distinct driver aggression profiles. This diversity is essential for training models capable of generalizing to the unconstrained realities of mixed traffic environments. The raw trajectory data was meticulously processed to extract valid lane-change maneuvers and corresponding lane-keeping sequences. A lane change is defined as the entire sequence of events starting from the initial lateral deviation to the moment the vehicle establishes a stable position in the target lane. To prevent the model from trivially predicting the maneuver after it has physically occurred, the observation window was strictly truncated prior to the vehicle crossing the lane markings. Extensive filtering was applied to remove anomalies, sensor noise, and highly erratic trajectories that resulted from recording errors.

Table 3: Characteristics of the Processed Evaluation Dataset

Dataset Metric	Total Recorded Instances	Left Lane Changes	Right Lane Changes	Lane Keeping
Highway Segment Alpha	14500	2300	2150	10050
Urban Expressway Beta	18200	3100	2800	12300
Merging Zone Gamma	9400	1850	1100	6450
Combined Aggregate	42100	7250	6050	28800

5.2 Implementation and Evaluation Metrics

The interaction field was discretized into a spatial grid spanning sixty meters longitudinally and twenty meters laterally around the ego vehicle, with a resolution of zero point five meters per cell. This resolution was chosen to balance computational efficiency with the need for fine-grained spatial modeling. The neural network components were implemented using standard deep learning libraries and trained utilizing stochastic gradient descent with momentum. A dynamic learning rate scheduler was employed, alongside early stopping mechanisms based on the validation loss to prevent overfitting. The hyper-parameters governing the interaction field decay formulas were initialized using domain knowledge from traffic flow theory and subsequently fine-tuned via backpropagation. To provide a comprehensive assessment of the model's performance, a diverse suite of evaluation metrics was utilized. Standard classification metrics, including Accuracy, Precision, Recall, and the F1-score, were employed to measure the deterministic predictive capabilities. Because lane-keeping instances heavily outnumber lane-change maneuvers, macro-averaged F1-scores were prioritized to ensure the model did not simply exploit the class imbalance by always predicting straight driving. Furthermore, to evaluate the efficacy of the uncertainty calibration module, Expected Calibration Error and Maximum Calibration Error were calculated. These metrics quantify the discrepancy between the predicted probabilities and the actual empirical frequencies across multiple probability bins.

6. Results and Discussion

6.1 Quantitative Performance Evaluation

The quantitative analysis demonstrates the unequivocal superiority of the Interaction-Field Residual Learning framework over established baselines. When compared against standard recurrent

ISSN: 2519-9439

architectures and graph-based models, the proposed method achieves significantly higher F1-scores across all maneuver categories. The performance gains are particularly pronounced in the early prediction of lane changes, typically one to three seconds prior to the actual lane boundary crossing. Standard models often exhibit severe latency in their predictions, only triggering a positive intent classification when the lateral movement becomes overwhelmingly obvious. In contrast, our model successfully detects the subtle, preliminary shifts in the interaction field, providing autonomous planning systems with crucial additional reaction time.

Table 4Quantitative Results for Lane-Change Prediction Intent

Evaluated Framework	Macro F1-Score	Precision (Left LC)	Recall (Right LC)	Expected Calibration Error
Standard Recurrent Network	0.76	0.78	0.72	0.145
Graph Convolutional Model	0.81	0.82	0.79	0.112
Attention-Based Transformer	0.84	0.86	0.81	0.098
Proposed Interaction-Field	0.91	0.92	0.89	0.034

A critical examination of Table 4 reveals the most significant advantage of the proposed framework: the drastic reduction in the Expected Calibration Error. While advanced transformer models provide competitive raw accuracy, their uncalibrated nature leads to an Expected Calibration Error nearly three times higher than our proposed approach. This indicates that the baseline models frequently produce highly confident incorrect predictions, a characteristic that is fundamentally incompatible with the stringent safety requirements of autonomous driving. By effectively minimizing the calibration error, our method ensures that the autonomous vehicle's decision-making algorithms can trust the probabilistic outputs, executing defensive maneuvers only when mathematically justified.

6.2 Qualitative Analysis of Mixed Traffic Scenarios

Beyond the aggregate statistical improvements, qualitative analysis of specific complex traffic scenarios reveals the deep contextual understanding achieved by the Interaction-Field Residual Learning model. Consider a densely packed highway scenario where the target vehicle is boxed in by slower-moving trucks. Standard models, relying primarily on distance metrics, often assign equal probability to left and right lane changes if the physical gaps appear geometrically similar. However, our interaction field dynamically models the velocity differentials. If the left lane is occupied by a fast-approaching vehicle, the resulting repulsive potential wave propagated forward through the field immediately suppresses the probability of a left lane change, mirroring the intuitive caution of a human driver. Furthermore, the separation of kinematics and residuals proves highly effective in anomaly handling. In situations where a leading vehicle performs an erratic, unexpected hard braking maneuver, the kinematic module correctly projects a collision trajectory based on pure physics. Simultaneously, the interaction-residual module processes the rapidly intensifying repulsive field, immediately recognizing the target vehicle's urgent intent to swerve or change lanes to avoid the collision. Models that entangle these two streams often suffer from delayed reactions, as the sudden shift in behavior contradicts the previously learned smooth temporal sequences. The distinct architectural boundary in our framework ensures that abrupt, high-intensity interaction signals can instantly override steady-state kinematic predictions.

ISSN: 2519-9439

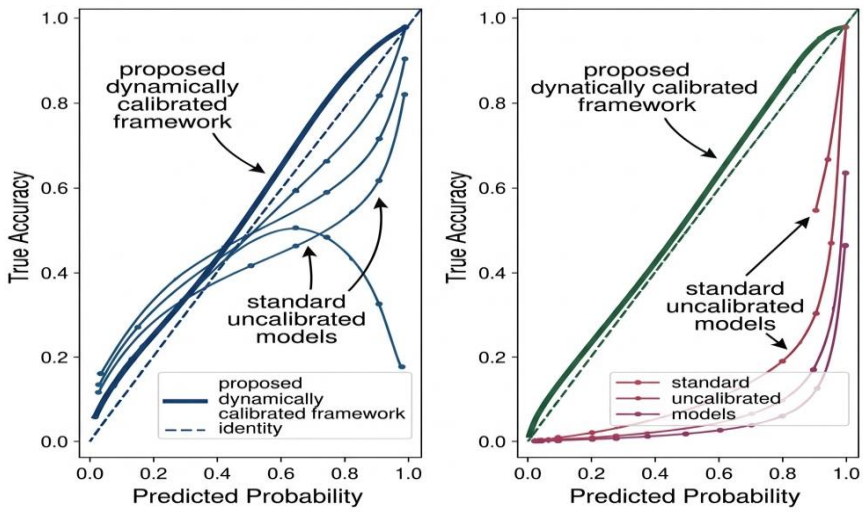


Figure 3Calibration Reliability Diagrams

7. Conclusion

7.1 Summary of Findings

This comprehensive study introduced the Interaction-Field Residual Learning framework, a paradigm-shifting approach for predicting lane-change intentions in complex, mixed traffic environments. By moving away from discrete, simplistic representations of surrounding vehicles, we successfully modeled the traffic environment as a continuous, dynamic field of interacting potentials. This continuous representation proved highly adept at capturing the non-linear, asymmetric influences that govern human driving behavior. The integration of a residual learning architecture allowed the neural network to isolate and learn these complex social behaviors without the confounding noise of standard vehicle kinematics. Most importantly, the implementation of a dynamic temperature scaling calibration module bridged the gap between theoretical accuracy and practical reliability, ensuring that the model's output probabilities genuinely reflect real-world maneuver likelihoods.

7.2 Limitations and Future Directions

Despite the significant advancements demonstrated, the current framework exhibits certain limitations that warrant further investigation. The interaction field computation, while mathematically elegant, introduces considerable computational overhead when scaling to highly congested urban intersections with dozens of closely interacting participants. The resolution of the spatial grid must be carefully optimized to prevent excessive memory consumption during real-time deployment on edge computing devices. Additionally, the current model assumes a relatively homogeneous distribution of human driver types. In reality, mixed traffic features a spectrum of aggressive and conservative driving styles, which could be better modeled by incorporating personalized latent variables into the interaction field equations. Future research will focus on developing computationally lighter variations of the field representation using sparse tensor operations, and exploring meta-learning techniques to rapidly adapt the interaction field parameters to the localized driving culture of different geographical regions. Integrating these enhancements will further solidify the foundation for safe, socially intelligent autonomous navigation in the coming era of mixed mobility.

References

Li, B., Zhang, R., Liang, H., Zhang, J., Zhang, J., Chen, X., ... & Wang, J. (2026). Interagent: Physics-based multi-agent command execution via diffusion on interaction graphs. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 15253-15265).

Wang, X., Ma, X., Zhu, P., Ng, W. S., Zhang, H., Xia, X., ... & Au, K. W. S. (2024). Design, optimization, and experimental validation of a handheld nonconstant-curvature hybrid-structure robotic instrument for maxillary sinus surgery. IEEE/ASME Transactions on Mechatronics, 29(4), 3074-3082.

Wang, J., Malawade, A. V., Zhou, J., Yu, S. Y., & Al Faruque, M. A. (2024, January). Rs2g: Data-driven scene-graph extraction and embedding for robust autonomous perception and scenario understanding. In 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (pp. 7478-7487). IEEE.

Zhao, J., Zhang, C., Wang, C., & Yang, P. (2025). Learning from expert factors: Trajectory-level reward shaping for formulaic alpha mining. arXiv preprint arXiv:2507.20263.

Prescribed-time tracking control of MIMO nonlinear systems with nonvanishing uncertainties

Xing, Z. (2026). A Synthesizable Bandwidth Monitoring Module for DDR4 Memory Controllers: RTL Design and Post-Implementation Analysis. Microelectronics Journal, 107410.

Yao, J., Yao, C., Chen, Y., Liu, X., Huang, D., & Chu, W. (2026, April). Explainable AI Enhanced Traffic Forecasting for Proactive Autonomous Vehicle Routing. In 2026 IEEE Conference on Technologies for Sustainability (SusTech) (pp. 1-7). IEEE.

Wang, Y., & Ling, C. (2025). Controlling attributes of. xpt files generated by R. In PharmaSUG 2025 conference proceedings. San Diego, CA.

Cheng, Xiaoyuan, et al. "How Does the Lagrangian Guide Safe Reinforcement Learning through Diffusion Models?." arXiv preprint arXiv:2602.02924 (2026).

Yang, Y., Wang, Q., Li, C., Hu, H., Wu, C., Jiang, Y., ... & Xu, B. (2025, May). Fewer may be better: Enhancing offline reinforcement learning with reduced dataset. In International Conference on Learning Representations (Vol. 2025, pp. 7076-7098).

Yang, Y., Hu, H., Li, W., Li, S., Yang, J., Zhao, Q., & Zhang, C. (2023, June). Flow to control: Offline reinforcement learning with lossless primitive discovery. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 37, No. 9, pp. 10843-10851).

Yang, Y., Ma, X., Li, C., Zheng, Z., Zhang, Q., Huang, G., ... & Zhao, Q. (2021). Believe what you see: Implicit constraint approach for offline multi-agent reinforcement learning. Advances in Neural Information Processing Systems, 34, 10299-10312.

Li, Y., Du, P., Ren, H., Xi, T., Gao, K., Kang, Z., ... & Zhang, Y. (2026). Mobile sensing discovery of when where and why vulnerable road users break traffic rules.npj Sustainable Mobility and Transport,3(1), 1.

Yu, Jiangbo, Raphaël Frank, Luis Miranda-Moreno, Sasan Jafarnejad, Jônatas Augusto Manzolli, Jiyao Wang, and Ali Eslami. Agentic Vehicles for Human-Centered Mobility: Definition, Prospects, and System Implications. 2026.

Wang, H., Li, Z., Zhang, Y., He, Z., Jiang, L., Li, K., ... & Gu, D. (2026). ConsistNav: Closing the Action Consistency Gap in Zero-Shot Object Navigation with Semantic Executive Control. arXiv preprint arXiv:2605.09869.

Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., et al. (2016). End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316.

LaValle, S. M. (1998). Rapidly-exploring random trees: A new tool for path planning. Technical Report, Iowa State University.

Kavraki, L. E., Švestka, P., Latombe, J.-C., & Overmars, M. H. (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. IEEE Transactions on Robotics and Automation, 12(4), 566–580.

Mur-Artal, R., & Tardós, J. D. (2017). ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Transactions on Robotics, 33(5), 1255–1262.

Driess, D., Xia, F., Sajjadi, M. S. M., Lynch, C., Chowdhery, A., Ichter, B., et al. (2023). PaLM-E: An embodied multimodal language model. In Proceedings of ICML.

ISSN: 2519-9439

Levine, S., Pastor, P., Krizhevsky, A., & Quillen, D. (2018). Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. International Journal of Robotics Research, 37(4–5), 421–436.

Ross, S., Gordon, G., & Bagnell, D. (2011). A reduction of imitation learning and structured prediction to no-regret online learning. In Proceedings of AISTATS.

Schulman, J., Levine, S., Abbeel, P., Jordan, M., & Moritz, P. (2015). Trust region policy optimization. In Proceedings of ICML.

Kalashnikov, D., Irpan, A., Pastor, P., Ibarz, J., Herzog, A., Jang, E., et al. (2018). QT-Opt: Scalable deep reinforcement learning for vision-based robotic manipulation. arXiv preprint arXiv:1806.10293.

Hutchinson, S., Hager, G. D., & Corke, P. I. (1996). A tutorial on visual servo control. IEEE Transactions on Robotics and Automation, 12(5), 651–670.

Li, Q., Ye, Q., Zhang, N., Zhang, W., & Hu, F. (2025). Digital-twin-enabled industrial IoT: Vision, framework, and future directions. IEEE Wireless Communications, 32(6), 173-181.

Liu, X., Yu, Z., Yao, Y., Yang, K., Zeng, H., & Chen, C. P. (2026). A Comprehensive Survey of Cross-Modal Retrieval: Breaking Through Modal Barriers. ACM Computing Surveys.

Lin, C. C., Li, G., Hung, C. C., Kuo, Y. C., & Peng, W. C. (2026). Rethinking suspicious account identification from a perspective of machine learning ranking. International Journal of Machine Learning and Cybernetics, 17(3), 84.

Sha, L., Wang, Y., Meng, P., Deng, Y., Chen, T., Zhang, X., ... & Xu, Q. (2025). Pharmacological inhibition of PSPH reduces serine levels and epileptic seizures. Nature Chemical Biology, 21(11), 1742-1753.

Yang, Z., Ji, W., Guo, Q., & Wang, Z. (2023, October). Javp: Joint-aware video processing with edge-cloud collaboration for dnn inference. In Proceedings of the 31st ACM International Conference on Multimedia (pp. 9152-9160).

Peng, Y., & Li, H. (2023). Benthic foraminiferal morphogroups at the end-Guadalupian extinction in eastern Sichuan Basin, China. Palaeogeography, Palaeoclimatology, Palaeoecology, 618, 111522.

Zhang, Y., Huang, Z., Zhao, M., Zhang, C., Lu, Y., Ji, Y., ... & Zeng, A. (2025). Learning unbiased cluster descriptors for interpretable imbalanced concept drift detection. IEEE Transactions on Emerging Topics in Computational Intelligence.

Jia, Y., Chen, Q., Mu, W., & Zhang, W. (2023). Managing the value co-creation of peer service providers in the sharing economy: the perspective of customer incivility. Heliyon, 9(6).

Zhu, R., Jiang, B., Mei, L., Yang, F., Wang, L., Gao, H., ... & Zhang, D. (2025). Adaptflow: Adaptive workflow optimization via meta-learning. arXiv preprint arXiv:2508.08053.

Kingma, D. P., & Welling, M. (2014). Auto-encoding variational Bayes. In Proceedings of ICLR.