A Conflict-Aware and Dynamically Adaptive Trajectory Prediction Method for Multi-Agent Traffic

Chen Wang[1,2,3], Haojie Li[1,2,3,*], Yuwen Zhang[1,2,3], Gang Ren[1,2,4]

[1]School of Transportation, Southeast University, NanJing, China
[2]Jiangsu Key Laboratory of Urban ITS, NanJing, China
[3]Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, NanJing, China

*Corresponding Author Email: h.li@seu.edu.cn, Tel: +86 13645172527

Abstract. In complex multi-agent traffic environments, the accuracy and safety awareness of trajectory prediction directly affect the reliability of decision-making in autonomous driving systems. Existing approaches primarily focus on fitting historical motion patterns, while lacking effective mechanisms for modeling interactive conflicts and risk-inducing behaviors, making it difficult to achieve stable prediction in high-density scenarios. To address these limitations, this paper proposes a conflict-aware and dynamically adaptable trajectory prediction framework for multi-agent traffic scenes. First, a two-dimensional time-to-collision (2D-TTC) matching mechanism is developed to jointly evaluate longitudinal and lateral TTC values, enabling the selective identification of influential interacting agents and suppressing irrelevant interaction noise at the source. Second, a dynamic model-switching strategy based on global interaction intensity is introduced to adaptively adjust network complexity according to traffic density and risk levels, thereby balancing prediction accuracy and computational efficiency. Furthermore, a conflict-aware weighted loss function is incorporated during training to enhance the model's sensitivity to high-risk interaction samples at the parameter-learning level, improving its ability to recognize potentially dangerous behaviors. Experiments conducted on the Waymo Open Dataset demonstrate that the proposed method achieves reductions in Average Displacement Error (ADE) and Final Displacement Error (FDE) compared with representative baseline models, while simultaneously improving inference efficiency. In addition, the method exhibits stronger risk perception and earlier conflict recognition under the post-hoc safety metric Alert Rate (AR). These results indicate that the proposed framework significantly enhances safety awareness and scene adaptability without compromising prediction accuracy, offering new insights and practical engineering value for safety-critical modeling and risk-aware trajectory prediction in intelligent transportation and autonomous driving systems.

Keywords: Multi-Agent Interaction, Dynamic Model Switching, Safety-Aware Prediction.

E. S. Zahran et al. (eds.), *Proceedings of the 2026 9th International Symposium on Traffic Transportation and Civil Architecture (ISTTCA 2026)*, Atlantis Highlights in Engineering 46,
https://doi.org/10.2991/978-94-6239-740-8_2

1 Introduction

Accurate trajectory prediction in complex multi-agent traffic environments is essential for safe decision-making in intelligent transportation systems and autonomous driving [1]. By inferring future motion from historical observations, trajectory predictors provide key inputs for planning and control. However, real-world traffic involves heterogeneous participants, dense interactions, and diverse behaviors, making traditional single-agent or purely kinematic models insufficient in highly dynamic scenes [2][3].

Deep learning has significantly advanced multi-agent trajectory prediction. Social-LSTM [4] captures temporal dependencies via RNNs, while Social-GAN [5], GATraj [6], and related attention/GNN-based methods model spatial interactions and achieve substantial improvements in ADE and FDE. Nevertheless, two major limitations remain:

(1) Lack of explicit conflict modeling—Most models treat all neighbors equally, causing irrelevant-interaction noise and reduced stability in dense traffic;

(2) Static architectural complexity: Fixed-capacity models cannot adapt to changing scene difficulty—leading to inefficiency in simple scenes and inadequate modeling in high-risk ones.

Moreover, mainstream evaluation focuses on accuracy metrics (ADE, FDE), while quantitative assessment of safety awareness—crucial for autonomous driving—remains largely unexplored.

To address these gaps, we propose a conflict-aware and dynamically adaptive trajectory prediction framework. The key components are:

(1) A 2D-TTC-based conflict matching mechanism that jointly evaluates longitudinal and lateral time-to-collision to identify high-impact neighbors and filter interaction noise.

(2) A dynamic model-switching strategy driven by global interaction intensity, enabling the predictor to adjust its complexity according to scene risk.

(3) A conflict-informed weighted loss, enhancing sensitivity to high-risk interactions and aligning predicted behavior with safety principles.

(4) A post-hoc safety metric (Alert Rate, AR) to evaluate early risk recognition.

Experiments on the Waymo Open Dataset [7] demonstrate that the proposed method reduces ADE by 2.3%, improves inference efficiency by 9.8%, and achieves substantially better conflict early-warning performance than strong baselines. These results show that the framework enhances both prediction accuracy and safety awareness, offering a practical path for risk-aware trajectory modeling in autonomous driving and intelligent transportation systems.

2 Literature Review

Trajectory prediction in multi-agent traffic environments is a central problem in intelligent transportation systems and autonomous driving. Early approaches—including Constant Velocity (CV) and Constant Acceleration (CA) models—rely on simplified motion assumptions [8] and thus fail to capture complex interactions or abrupt intention changes. With the development of deep learning, neural networks have increasingly been used to extract temporal dependencies and interaction cues automatically, significantly improving predictive capability.

RNN-based architectures (e.g., LSTM, GRU) have been widely employed to model the sequential nature of agent motion. Social-LSTM [4] introduced Social Pooling to capture spatially neighboring influences, followed by Social-GAN [5], which leverages adversarial learning for multimodal trajectory generation. More recent approaches such as GATraj [6] and Trajectron++ [9] utilize attention mechanisms and graph neural networks to encode relational structures among agents, achieving state-of-the-art performance in various datasets.

Despite these advances, existing models often lack explicit reasoning about potential conflicts. While several works incorporate risk heuristics such as time-to-collision (TTC) or collision probability [2], these components are typically appended as external modules rather than integrated into the trajectory generation pipeline. Furthermore, many interaction encoders treat all neighboring agents equally, introducing substantial irrelevant-noise in dense urban settings and weakening robustness.

Recent research has explored adaptive architectures to improve responsiveness to dynamic traffic environments. For instance, the Density-Adaptive Mobility Model (DAMM) [10] adjusts interaction ranges and influence probabilities according to scene density, while TRACER [11] employs adaptive interaction extraction with trajectory generation and auxiliary intention-recognition tasks. However, these methods generally rely on density or domain cues rather than conflict-aware triggers, limiting their precision in high-risk scenarios and creating insufficient synergy between safety perception and adaptability.

Most trajectory prediction studies still evaluate models primarily through accuracy metrics (ADE, FDE), overlooking the need for explicit safety awareness. As a result, models may achieve low positional error yet fail to detect high-risk interactions, which is problematic for safety-critical applications. The present work addresses this gap by integrating conflict-aware representation learning and a risk-sensitive weighted loss, emphasizing safety perception alongside predictive accuracy..

In summary, although recent trajectory prediction research has made progress in temporal modeling and accuracy enhancement, three major challenges remain:

(1) Lack of explicit conflict modeling—risk signals are seldom integrated into end-to-end prediction and rarely differentiate interaction intensity;

(2) Insufficient dynamic adaptability—Existing adaptive mechanisms primarily respond to density or domain-shift patterns rather than conflict-driven cues;

(3) Incomplete safety evaluation—current metrics emphasize accuracy but provide limited assessment of early conflict recognition.

This work tackles these challenges through a framework combining conflict-aware matching, dynamic structural adaptation, and safety-oriented loss design to enhance both prediction accuracy and risk perception.

3 Methodology

3.1 Problem Formulation

The goal of trajectory prediction is to infer the future motion states of a traffic agent based on its observed historical state sequence. For a target agent, let the observed sequence within an observation window of length t_{obse} be denoted as $X = \{p^1, p^2, \ldots, p^{t_{obse}}\}$, where each state $p^t = \{x^t, y^t, \ldots, a^t\}$ includes spatial coordinates, velocity, acceleration, and other dynamic attributes at time t. The prediction task aims to generate the future state sequence within a prediction horizon $Y = \{p^{t_{obse}+1}, p^{t_{obse}+2}, \ldots, p^{t_{obse}+t_{pred}}\}$.

To model multi-agent interactions, the input is further extended to include the collective state of all n agents in the scene at each time step. The scene state at time t is represented as a matrix ($p^t \in R^{n\times d}$, where d denotes the feature dimension of each agent). This formulation enables the model to capture the dynamic spatial–temporal dependencies and interaction relationships among multiple agents in traffic environments.

3.2 Two-Dimensional TTC (2D-TTC) Conflict Matching Mechanism

To address the limitation of implicit conflict modeling in existing trajectory prediction frameworks, this study introduces a two-dimensional time-to-collision (2D -TTC) matching mechanism to identify key interacting agents. The procedure is as follows:

Local Coordinate Transformation.

A target-centric coordinate frame is constructed using the agent's initial position and heading, ensuring interaction evaluation is independent of global map geometry.

Two-Dimensional TTC Computation.

Longitudinal and lateral TTC are computed as

$$TTC_x = \frac{x_B - x_A - l_A}{v_A^x - v_B^x}$$

$$TTC_y = \frac{y_B - y_A - w_A}{v_A^y - v_B^y}$$

$$TTC_{avg} = (|TTC_x| + |TTC_y|)/2$$

where x_A, y_A, x_B,y_B are the transformed coordinates; v_A^x, v_A^y, v_B^x, v_B^y are velocity components in the local frame; and l_A, w_A denote the length and width of the target agent.

Conflict Identification.

A neighbor is considered a key interacting agent if $TTC_{avg} < 3s$ at any point in the observation window. This removes irrelevant neighbors and reduces interaction noise.

3.3 Dynamic Model Switching Strategy Based on Global Interaction Intensity

To improve adaptability across scenes with different complexity levels, we adopt a dynamic model-switching strategy.

Global Interaction Intensity.

Scene-level complexity is quantified using the proportion of key interacting agents. The global interaction intensity I is defined as:

$$I = \frac{Number\ of\ conflict\ related\ agent\ pairs}{Total\ number\ of\ agent\ pairs}$$

Higher values of I indicate dense or highly interactive scenes (e.g., congested traffic), while lower values correspond to sparse or low-risk environments.

Model-Switching Rule.

Based on validation results across different I intervals, the system selects during inference the model that best matches the current scene intensity, enabling a balance between prediction accuracy and computational cost.

3.4 Dynamic Model Switching Strategy Based on Global Interaction Intensity

To increase sensitivity to high-risk interactions, a conflict-informed weighting term is incorporated into the loss:

$$\omega_{ij}^t = s_{ij} \cdot \boldsymbol{\sigma}(TTC_{ij}^{pred} - TTC_{ij}^{true})$$

where the severity index $s_{ij} \in [0,1]$ is computed by $SI = e^{-(\frac{TTC^2}{2PRT^2})}$[12], and $\boldsymbol{\sigma}(\cdot)$ is the Sigmoid activation. TTC_{ij}^{pred} and TTC_{ij}^{true} denote the predicted and ground-truth TTC values, respectively. Based on this weighting strategy, we design a regularization loss using a two-dimensional Gaussian distribution:

$$L_{reg} = -\frac{1}{N^2}\sum_{t=t_{obse}+1}^{t_{pred}}\sum_{i=1}^{N}\sum_{j=1}^{N}(1+\omega_{ij}^t) \cdot log\ (\boldsymbol{\mathcal{N}}(x_i^t, y_i^t \mid \mu_{1,i}^t, \mu_{2,i}^t, \sigma_{1,i}^t, \sigma_{2,i}^t, \rho_i^t))$$

where $\boldsymbol{\mathcal{N}}(\cdot)$ denotes the probability density of a bivariate Gaussian distribution; $\mu_{1,i}^t, \mu_{2,i}^t$ are mean position estimates; $\sigma_{1,i}^t, \sigma_{2,i}^t$ are standard deviations; and ρ_i^t is the correlation coefficient.

By assigning larger weights to conflict-prone samples, the model learns safer and more risk-aware trajectory distributions.

3.5 Overall Model Architecture

The proposed Conflict-Aware Multi-Agent Trajectory Prediction model (CM-TP) adopts an encoder–decoder architecture. The complete pipeline is illustrated in Figure 1 and consists of the following components:

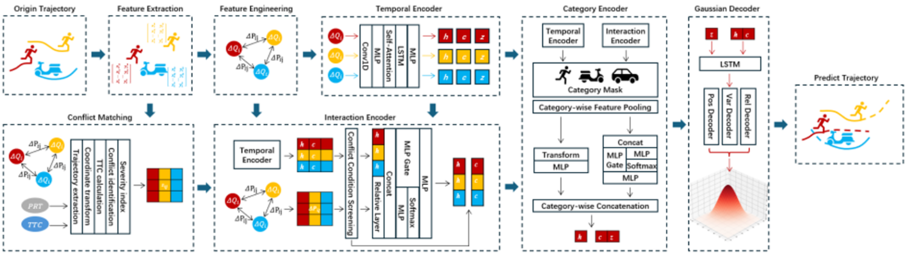

Fig. 1. Overall architecture of the proposed Conflict-Aware Multi-Agent Trajectory Prediction model (CM-TP)

As illustrated in Table 1, the CM-TP framework follows an encoder–decoder architecture that integrates temporal modeling, interaction reasoning, and semantic category encoding. Feature preprocessing reduces trajectory variance, while the temporal encoder extracts multi-scale motion patterns. The interaction encoder focuses on high-risk neighbors selected via 2D-TTC, reducing noise in dense traffic scenes. Category embeddings enhance behavior differentiation across agent types. Finally, a Gaussian decoding head generates uncertainty-aware future trajectories.

Table 1. Overview of CM-TP Model Architecture

Module	Description
Feature Engineering	Applies displacement normalization and relative-motion encoding to reduce scene dependency and highlight interaction cues.
Temporal Encoding	Combines 1D CNN + Transformer + LSTM to capture short-term motion patterns and long-range temporal dependencies.
Interaction Encoding	Utilizes 2D-TTC filtering to select key interacting agents, followed by attention-based fusion to generate interaction-aware features.
Category Encoding	Introduces agent-type embeddings (car, cyclist, pedestrian) to model heterogeneous motion behaviors.
Decoding & Prediction	Outputs a bivariate Gaussian distribution (μ, σ, ρ) for each future step, enabling probabilistic trajectory prediction.

4 Experimental Design

4.1 Dataset and Experimental Environment

4.1.1 Dataset Introduction. Experiments are conducted on the Waymo Open Motion Dataset (WOMD), a large-scale benchmark for urban multi-agent trajectory prediction. WOMD contains real-world driving data collected from multiple U.S. cities, with each scene lasting about 20 seconds at 10 Hz. The dataset includes heterogeneous traffic participants—vehicles, cyclists, and pedestrians—and captures dense interactions and diverse risk conditions, making it well suited for evaluating safety-aware forecasting models.

A total of 798 scenes are used for training, 202 for validation, and 150 for testing. Since portions of the official WOMD test trajectories are not publicly released, the validation set is used as the test set for reporting performance. To improve generalization, an internal split of the training set is adopted for hyper-parameter tuning and early stopping.

For preprocessing, the observation window is set to t_{obse} = 2 s (20 frames) and the prediction horizon to t_{pred} = 3 s (30 frames), along with additional 4-frame and 6-frame down-sampling for training efficiency. All trajectories are transformed into a target-centric local coordinate system defined by the agent's initial position and heading direction, removing map-dependent variations and stabilizing interaction representation.

4.1.2 Experimental Environment. All experiments are conducted on a workstation running Windows 11, equipped with an Intel Core i5-13490F CPU, NVIDIA GeForce RTX 4070Ti GPU (12 GB VRAM), and 32 GB RAM. The software environment consists of Python 3.10, PyTorch 2.1.0, and CUDA 12.2.

4.1.3 Model Implementation Details. The CM-TP model adopts a lightweight hierarchical encoding–decoding architecture. The temporal encoder uses a 1D convolution layer (kernel = 3, stride = 1, channels = 64) followed by a two-layer Transformer Encoder (hidden dim = 256, 8 heads) to capture both local motion trends and long-range dependencies.

The interaction encoder applies the proposed 2D-TTC–based conflict-matching mechanism to retain only safety-relevant neighbors, after which gated attention aggregation is performed to obtain a 256-dimensional interaction representation.

A category embedding module (dim = 32) models behavioral differences among cars, bicycles, and pedestrians, and pooled category-level features are fused with interaction features.

The decoder is a single-layer LSTM (hidden size = 256, 6 steps), which outputs bivariate Gaussian parameters for probabilistic trajectory prediction.

For training, Adam is used with an initial learning rate of 1×10^{-5}, batch size 64, and 300 epochs with Early Stopping. The TTC threshold in the conflict-matching module is selected based on preliminary calibration experiments. All experiments use fixed random seeds for reproducibility.

4.2 Baseline Models

To evaluate the effectiveness and generalizability of the proposed CM-TP framework, five representative trajectory prediction models—covering RNN-, GAN-, GNN-, and attention-based approaches—are selected as baselines. All models are re-trained using identical data splits, optimization settings, and evaluation metrics (ADE, FDE, AR) to ensure a fair comparison. A brief overview is provided in Table 2.

Table 2. Overview of Baseline Models

Model	Description
LSTM	Classical LSTM encoder–decoder capturing temporal dependencies in single-agent motion.
Social-LSTM[4]	Early interaction-aware model using Social Pooling to aggregate neighboring agents' features.
Social-GAN[5]	GAN-based multimodal predictor enhancing trajectory diversity via adversarial learning.
TrafficPredict[13]	4D spatiotemporal graph model capturing both category-level and instance-level correlations.
GATraj[6]	Graph Attention Network–based method focusing on key interactions in dense traffic scenes

These baselines span the dominant modeling paradigms in multi-agent trajectory forecasting and provide strong reference points for evaluating the CM-TP model.

4.3 Evaluation Metrics

To comprehensively evaluate both the prediction accuracy and safety-awareness capability of the proposed CM-TP model, three categories of metrics are adopted: Average Displacement Error (ADE), Final Displacement Error (FDE), and Alert Rate (AR). ADE and FDE assess trajectory prediction accuracy, while AR quantifies the model's ability to anticipate potential traffic conflicts.

Accuracy Metrics.

Accuracy is evaluated using ADE and FDE, defined as:

$$ADE = \frac{1}{N\times T}\sum_{i=1}^{N}\sum_{t=1}^{T}\|\hat{p}_i^t - p_i^t\|_2$$

$$FDE = \frac{1}{N}\sum_{i=1}^{N}\|\hat{p}_i^T - p_i^T\|_2$$

ADE measures average alignment over the prediction horizon, while FDE focuses on the final-step error, reflecting long-term prediction reliability.

Safety Metrics.

To measure the model's capability to foresee potential risks, we introduce the Alert Rate (AR), computed based on the Time-to-Collision (TTC) matrices.

For a scene containing N dynamic agents, TTC matrices for the observation phase, prediction phase, and ground-truth future evolution are denoted as TTC_{ij}^{obse}, TTC_{ij}^{pred} and TTC_{ij}^{true}, respectively. We define the indicator function:

$$\delta_{ij} = \begin{cases} 1, & if\ TTC_{ij}^{pred} < TTC_{ij}^{true} \\ 0, & otherwise \end{cases}$$

A positive indicator (δ_{ij}=1) implies that the model anticipates the conflict earlier than what is observed from the ground-truth future trajectory—providing the decision-making system with additional safety margin. (This corresponds to the distinction illustrated in Figure. 2, where two predictions with similar geometric accuracy may differ significantly in safety implications.)

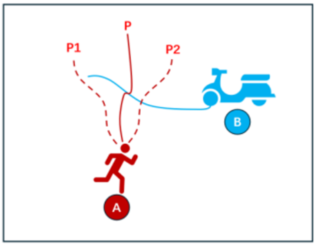

Fig. 2. Comparison of safety-awareness between different trajectory predictions.

To avoid counting irrelevant weak interactions, only the set $\mathcal{P}$ of agent pairs whose TTC_{ij}^{obse} is below a predefined threshold is considered. AR is then computed as:

$$AR = \frac{1}{|\mathcal{P}|} \sum_{(i,j) \in \mathcal{P}} \delta_{ij}$$

In this study, AR is evaluated under three TTC thresholds—6 s, 4 s, and 2 s—covering low-, medium-, and high-risk interaction levels.

Efficiency Metrics.

Real-time performance is measured using inference time:

$$E_{time} = \sum_{i=1}^{N} t_i^{end} - t_i^{start}$$

where t_i^{start} and t_i^{end} represent the start and end timestamps of the inference process for the i-th trajectory in the evaluation set.

4.4 Experimental Results and Analysis

4.4.1 Statistical Description of the Dataset. To support the subsequent evaluation of trajectory prediction models, we conduct a statistical analysis of the dataset constructed from the WOMD. Using the sliding-window segmentation strategy, continuous multi-agent trajectories are transformed into fixed-length samples. The resulting dataset follows a tensor format of $[N, T, H]$, where N is the total number of samples, T the number of time steps, and H the feature dimension.

Sample Statistics.

Table 3 summarizes the number of samples in the training, validation, and test sets.

Table 3. Sample Statistics

Dataset	Number of Samples
Training	13553
Validation	9001
Test	5695

Trajectory and Category Distribution.

To understand the composition of different traffic participants, Table 4 reports the number of trajectories corresponding to vehicles (Veh), bicycles (Bic), and pedestrians (Ped). These statistics reflect the diversity and imbalance inherent in urban traffic scenes.

Table 4. Trajectory Count and Category Distribution

Dataset	Number of Traj	Veh	Bic	Ped
Training	274382	329043	179787	156322
Validation	440088	222123	119237	98728
Test	274382	140787	71476	62119

4.4.2 Overall Performance Comparison. We compare the proposed CM-TP model with five representative baselines—LSTM, Social-LSTM (SL), Social-GAN (SG), TrafficPredict (TP), and GATraj—under identical data splits and training settings. Evaluation is conducted on the Waymo dataset using ADE, FDE, AR (under TTC < 6 s), and E_{time}. Results are summarized in Table 5.

Table 5. Performance Comparison on the Waymo Dataset

Model	ADE	FDE	AR	E_{time}
LSTM	3.7880	6.7014	0.4935	19.8441
SL	2.7786	4.9727	0.5134	75.6183
SG	2.4633	4.3829	0.5019	56.6678
TP	1.6288	2.9519	0.5211	120.1134
GATraj	0.8602	1.4996	0.5236	104.8178
CM-TP	0.8408	1.5086	0.6482	95.4253

CM-TP achieves accuracy on par with the strongest baselines (e.g., GATraj), while delivering a substantial improvement in safety awareness, evidenced by the highest AR score. The gain comes from the conflict-aware matching and risk-weighted probabilistic modeling introduced in our framework. Although its inference time is moderately higher than LSTM-based models, CM-TP remains computationally comparable to modern attention- and GNN-based approaches.

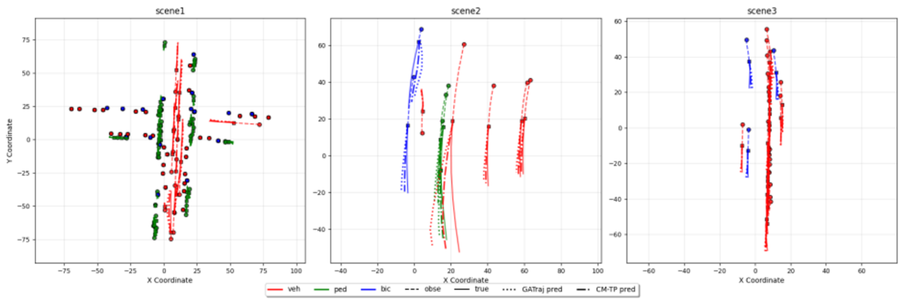

Fig. 3. Visualization of trajectory prediction results in selected scenarios

Figure 3 illustrates qualitative comparisons in three representative scenarios. CM-TP produces trajectories that are not only geometrically accurate but also more conservative and stable in high-risk or highly interactive scenes, highlighting its improved safety sensitivity.

4.4.3 Ablation Study. To examine the independent contribution and joint effect of each core component in the proposed CM-TP framework, we conduct a comprehensive ablation study. Specifically, we remove or replace the 2D-TTC conflict matching module, the interaction modeling layer, and the distribution-weighted loss to analyze their influence on prediction accuracy, safety sensitivity, and inference efficiency. All variants are trained under identical configurations to ensure fair comparison. Table 6 summarizes the structural definitions of the ablated models:

Table 6. Definitions of Ablation Model Variants

Model	Structural Description
ST-TP	Temporal encoder + category embedding
CT-TP	Temporal encoder + interaction encoder + category embedding
CM-TP	Temporal encoder + 2D-TTC conflict matching + interaction encoder + category embedding
SI-TP	CM-TP + SI-weighted loss
TTC-TP	CM-TP + TTC-difference–weighted loss

To facilitate comparison, ADE, AR, and E_{time} are selected as the primary evaluation metrics. Table 7 reports the performance of all ablated models under the full testing set and three risk thresholds (TTC < 6 s, 4 s, 2 s).

Table 7. Performance Comparison Across Ablation Settings

Model	ADE	AR Ped			E_{time}
		$TTC < 6$	$TTC < 4$	$TTC < 2$	
ST-TP	1.1016	0.6125	0.6513	0.7070	**82.5952**
CT-TP	0.8727	0.6719	0.6958	0.7777	112.9965
CM-TP	0.8408	0.6482	0.6972	0.7575	95.4253
SI-TP	**0.7779**	0.6971	0.7647	0.8080	97.1324
TTC-TP	0.8661	**0.7376**	**0.7819**	**0.8787**	96.8676

The results reveal several insights:

(1) Interaction modeling (CT-TP) significantly improves ADE over the pure temporal model (ST-TP), confirming the importance of capturing multi-agent dependencies.

(2) Conflict matching (CM-TP) enhances prediction in high-risk conditions (TTC < 4 s), verifying the necessity of noise filtering and conflict-aware neighbor selection.

(3) Risk-weighted losses (SI-TP, TTC-TP) further strengthen safety sensitivity, with TTC-TP achieving the highest AR across all thresholds, highlighting the effectiveness of leveraging TTC discrepancy during training.

These findings collectively demonstrate that the proposed components contribute to both accuracy and risk awareness, and their combination yields the best overall performance.

4.4.4 Performance Comparison Across Different Risk Levels. To systematically evaluate model behavior under varying traffic risk conditions, the conflict-rate interval [0%, 100%] is divided into ten equal-risk bins, each representing an increasingly challenging interaction setting—from low-risk sparse traffic to highly interactive and collision-prone scenarios.

For each risk bin, we compute the average ADE for all baseline models (LSTM, SL, SG, TP, GATraj), the ablation models (ST-TP, CT-TP, CM-TP), and the final proposed model (CM-TP (Proposed)). The performance trends are visualized as bar charts, as shown in Figure 4.

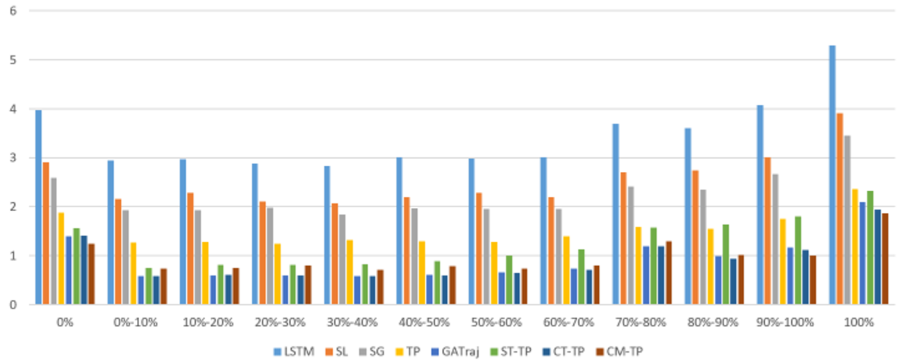

Fig. 4. Performance comparison under different risk-level scenarios

4.4.5 Performance of Dynamic Model Switching. Based on the observations in Section 4.4.4, we determine the optimal model-switching configuration for the given dataset. To further examine the adaptability of multi-model collaboration in dynamic scenarios, we design a conflict-rate–based switching experiment.

We select long-horizon trajectory sequences (>19 s) from the test set, resulting in a total of 2359 usable samples. These scenarios feature continuously evolving traffic density, interaction complexity, and conflict rates, making them ideal for evaluating dynamic model selection. So each sample is predicted under two settings:

(1) Single-model prediction (no switching)

(2) Dynamic switching, where the model is selected based on the current risk level

We compute the cumulative ADE and FDE over the full trajectory horizon. The results are summarized in Table 8.

Table 8. Cumulative Prediction Errors with and Without Dynamic Model Switching

Model	ADE	FDE
LSTM	60.1584	103.1845

Model	ADE	FDE
SL	40.1278	75.5694
SG	38.1204	66.9565
TP	26.8674	51.5852
GATraj	16.6451	30.6396
ST-TP	19.1478	36.7726
CT-TP	16.4434	30.0257
CM-TP	16.8458	31.1135
Mix-Model	**15.9508**	**29.0564**

The dynamic switching model ("Mix-Model") achieves the lowest cumulative ADE and FDE across the long-horizon sequences, outperforming all single-model baselines. The improvement is particularly notable in high-risk segments, where matching model complexity to interaction intensity allows more accurate trajectory adaptation. These findings highlight the advantage of combining conflict-aware filtering and dynamic structural adjustment, confirming the practical feasibility of multi-model cooperative prediction in real-world autonomous driving scenarios.

5 Discussion and Conclusion

5.1 Research Summary

This study addresses multi-agent trajectory prediction in urban traffic by introducing a conflict-aware and dynamically adaptive framework. A 2D-TTC–based matching mechanism is proposed to identify truly influential neighbors, effectively reducing interaction noise. A risk-weighted probabilistic loss further incorporates conflict severity into the learning process, enabling the model to emphasize high-risk samples and generate safety-consistent trajectory distributions. Additionally, a dynamic model-switching strategy adjusts prediction complexity according to scene-level interaction intensity, balancing accuracy, robustness, and efficiency.

Experiments on the Waymo Open Motion Dataset demonstrate that CM-TP outperforms representative baselines (LSTM, Social-LSTM, Social-GAN, TrafficPredict, GATraj) on ADE/FDE while also achieving stronger early conflict recognition under the AR safety metric. Moreover, the proposed switching mechanism stabilizes performance in long-horizon sequences. Overall, CM-TP integrates interaction filtering, risk-aware optimization, and structural adaptivity into a unified solution, advancing both prediction accuracy and safety sensitivity for autonomous driving.

5.2 Engineering Significance and Future Work

The proposed conflict-aware and adaptive prediction framework has practical engineering value in real traffic applications:

(1) Improved Risk Anticipation——by combining 2D-TTC filtering with risk-weighted distribution modeling, the system offers earlier detection of hazardous interactions, benefiting downstream planning and emergency intervention modules.

(2) Adaptive Real-Time Deployment——the dynamic switching mechanism allocates computational resources according to scene complexity, enabling efficient operation in both low-risk and high-density environments—critical for onboard autonomous driving systems.

(3) Lightweight V2X Compatibility——since the conflict-matching module relies only on geometric and kinematic inputs, it can be easily integrated into vehicle–infrastructure cooperation systems with limited bandwidth and sensing capability.

Future work may further enhance the framework along several directions:

(1) Richer Risk Modeling——incorporating acceleration anomalies, intention inconsistency, and safety zones to build a more comprehensive interaction-risk descriptor.

(2) Safety-Aware Learning Objectives——embedding AR-like safety indicators directly into supervised loss functions or reinforcement learning rewards to jointly optimize accuracy and safety performance.

References

1. Xi S, Liu Y, Fan Q, et al. HAVING IT ALL: Multi-Agents, Accuracy and Interpretability in Trajectory Prediction for Autonomous Driving[J]. Authorea Preprints, 2024. https://www.techrxiv.org/doi/10.36227/techrxiv.173152254.49960816
2. Fang J, Zhu C, Zhang P, et al. Heterogeneous trajectory forecasting via risk and scene graph learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(11): 12078-12091. https://doi.org/10.1109/TITS.2023.3287186
3. Choi C H, Park S W, Park J, et al. Prediction and prevention of crowd-crush accidents using crowd-density simulation based on unity engine[J]. Discover Applied Sciences, 2024, 7(1): 33. https://doi.org/10.1007/s42452-024-05962-8
4. Alahi A, Goel K, Ramanathan V, et al. Social lstm: Human trajectory prediction in crowded spaces[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 961-971. https://doi.org/10.1109/CVPR.2016.110
5. Gupta A, Johnson J, Fei-Fei L, et al. Social gan: Socially acceptable trajectories with generative adversarial networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 2255-2264. https://doi.org/10.1109/CVPR.2018.245
6. Cheng H, Liu M, Chen L, et al. Gatraj: A graph-and attention-based multi-agent trajectory prediction model[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2023, 205: 163-175. https://doi.org/10.1016/j.isprsjprs.2023.09.013
7. Ettinger S, Cheng S, Caine B, et al. Large scale interactive motion forecasting for autonomous driving: The waymo open motion dataset[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 9710-9719. https://doi.org/10.1109/ICCV48922.2021.00959
8. Wang Y, Wang J, Jiang J, et al. SA-LSTM: A trajectory prediction model for complex off-road multi-agent systems considering situation awareness based on risk field[J]. IEEE Transactions on Vehicular Technology, 2023, 72(11): 14016-14027. https://doi.org/10.1109/TVT.2023.3287207

9. Westerhout F S B, Schumann J F, Zgonnikov A. Smooth-Trajectron++: Augmenting the Trajectron++ behaviour prediction model with smooth attention[C]//2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2023: 5423-5428. https://doi.org/10.1109/ITSC57777.2023.104227
10. Wen D, Xu H, He Z, et al. Density-adaptive model based on motif matrix for multi-agent trajectory prediction[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 14822-14832. https://doi.org/10.1109/CVPR52733.2024.01422
11. Qian H, Shi X, Hawbani A, et al. TRACER: Transfer Knowledge-Based Collaborative Vehicle Trajectory Prediction for Highway Traffic Toward Cross-Region Adaptivity[J]. IEEE Transactions on Intelligent Transportation Systems, 2025. https://doi.org/10.1109/TITS.2025.3497863
12. Tageldin A, Sayed T, Ismail K. Evaluating the safety and operational impacts of left-turn bay extension at signalized intersections using automated video analysis[J]. Accident Analysis & Prevention, 2018, 120: 13-27. https://doi.org/10.1016/j.aap.2018.07.021
13. Ma Y, Zhu X, Zhang S, et al. Trafficpredict: Trajectory prediction for heterogeneous traffic-agents[C]//Proceedings of the AAAI conference on artificial intelligence. 2019, 33(01): 6120-6127. https://doi.org/10.1609/aaai.v33i01.4569