IMTP: Intention-Matching Trajectory Prediction for Autonomous Vehicles

Wenzhi Bai, Luwen Yu, Andrew Weightman, Zhengtao Ding, Zhiqiang Zhang,
Shengquan Xie, Zhenhong Li

***Abstract*—Trajectory prediction for surrounding vehicles is critical for ensuring the safety of autonomous driving. In this paper, we introduce a novel prediction framework named Intention-Matching Trajectory Prediction (IMTP). Different from existing results that predict trajectories based on only environmental information and historical trajectories, the proposed method initially identifies the possible intentions of surrounding vehicles based on the environment and generates intention-informed trajectories based on the physical vehicle model. Historical trajectories are then used to identify the intention and trajectory with the highest probability. The proposed framework effectively integrates the physical vehicle model, road-related environmental factors, and interactions among surrounding vehicles. A comparative study conducted on a public dataset demonstrates that our framework enhances both prediction accuracy and robustness.**

***Index Terms*—trajectory prediction, autonomous vehicles**

I. Introduction

Autonomous driving is one of the most exciting areas in the last decades due to its potential to transform travel modalities and offer safer, more comfortable, and more efficient journeys [1]. While the integration of lower-level autonomy into driver-assistance systems has yielded significant progress, the realization of high-level autonomous driving still presents considerable challenges. One of the major difficulties is accurately predicting the trajectories of surrounding traffic participants, which is crucial to enable intelligent downstream decision-making and planning [2], [3].

Physics-based approaches [4]–[6] employ rules and physical models [7]. Their prediction accuracy relies on modelling the behaviour of surrounding vehicles which is generally complex. In contrast, learning-based approaches [8]–[10] can give accurate intention predictions due to their advantage in modelling implicit driving behaviour. Nonetheless, due to the absence of constraints from rules and physical models, learning-based approaches may generate forecasting trajectories incompliant with kinematic or environmental characteristics. In recent years, there has been a growing number of studies combining the strengths of the two, focusing on more reliable long-term trajectory predictions [11]–[13].

Experienced human drivers are capable of predicting the trajectories of surrounding vehicles intuitively and instantaneously. During the process, the human driver will initially analyze the possible intentions of surrounding vehicles based on their historical trajectories and environment (e.g., lane connectivity, obstacles and free space), and generate intention-related trajectories. Then, the driver will continually verify and update these intentions and trajectories based on new observations. The process significantly differs from the existing approach, which predicts trajectories without considering the intentions of surrounding vehicles.

Inspired by human drivers, this paper proposes a novel Intention-Matching Trajectory Prediction (IMTP) framework composed of a physics-based trajectory generator, a learning-based trajectory matcher, and a Kalman-based intention estimator. In this framework, we first identify potential target states after a few seconds, including target reference paths, target speeds and etc., based on road connectivity and vehicle kinematic constraints in the environment. These potential target states are regarded as intentions. Planning trajectories are then generated iteratively from the initial moment to the final moment of historical observation, with each state at the iterative moment as the start state and the identified target state as the end state. The observed trajectories between the planning starting moment and the final moment are matched with the planned trajectories, resulting in the computation of probabilities for each trajectory. This iterative process continues until the final historical observation moment or the present moment, producing probabilities for all intentions and their predicted trajectories. Subsequently, a subset of trajectories with the highest probabilities is chosen as outputs, catering to specific output requirements. Thus, the intention acts as an intermediate step to obtain the best predicted trajectory by matching observable behaviour with implicit intentions.

This framework effectively combines explicit vehicle and environmental constraints with implicit intentions, enhancing the reliability of trajectory predictions. Comparative studies, conducted using the open dataset Argoverse [14], demonstrate the accuracy and robustness of the proposed approach.

II. Problem formulation

In this section, the problem of trajectory prediction is formulated and the proposed IMTP framework is introduced.

A. Formulation of Trajectory Prediction

In the context of trajectory prediction, we only consider a selected target vehicle a_{tar} among on-road vehicles $\mathcal{A}$. For simplicity, all subsequent marks will not be explicitly denoted as the target vehicle. Let us consider that a self-driving vehicle is equipped with detection and tracking modules to measure the states x^t of a target vehicle at discrete time instants $t \in \{t_0, t_1, \cdots, t_k\}$ with a constant frequency and has access to high-definition (HD) map $\mathcal{M}$, where x includes positions, velocities, accelerations, heading and angular velocity in 2D and $\mathcal{M}$ includes environmental information, i.e., the boundaries, centerline and traffic direction of drive lanes. The objective of trajectory prediction in this paper is to predict multi-modal future trajectories with corresponding probabilities of the target vehicle $\mathcal{R}^{t_k} = \{(\mathcal{T}_n^{t_k}, \mathcal{P}_n^{t_k})|n = 1, 2, \cdots, N\}$ from the final observed moment t_k subject to constraints $\mathcal{C} = \mathcal{C}_\mathcal{M} \cup \mathcal{C}_\mathcal{V}$, where N is the number of predicted trajectories, $\mathcal{T}_n^{t_k} = \{x_n^t|t = t_{k+1}, \cdots, t_k + \tau_{pre}\}$ is the nth predicted trajectory with continuous state information up to the prediction horizon τ_{pre}, $\mathcal{P}_n^{t_k}$ is the corresponding trajectory probability of $\mathcal{T}_n^{t_k}$, $\mathcal{C}_\mathcal{M}$ are environmental constraints derived from HD map, and $\mathcal{C}_\mathcal{V}$ are kinematic constraints of the target vehicle. We also use $\mathcal{X}^{t_0} = \{x^t|t = t_0, \cdots, t_k\}$ to represent all observed states x from the initial observed moment t_0.

B. The Framework of IMTP

Our method chooses a three-stage architecture, as shown in Fig. 1, which is like a common two-stage architecture with intermediate results. The framework of IMTP consists of a physics-based trajectory generator G, a learning-based trajectory matcher M, and a Kalman-based intention estimator E.

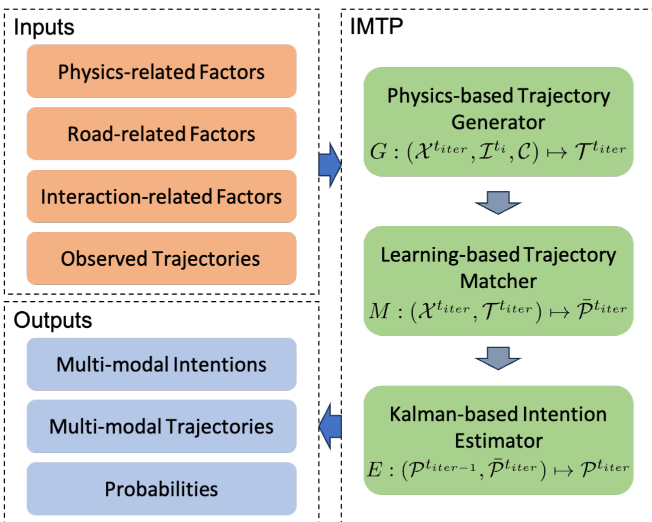


Fig. 1: The Framework of IMTP

a) Physics-based trajectory generator: we use the generator $G : (\mathcal{X}^{t_{iter}}, \mathcal{I}^{t_i}, \mathcal{C}) \mapsto \mathcal{T}^{t_{iter}}$ to produce corresponding reachable trajectories $\mathcal{T}^{t_{iter}}$ from $\mathcal{X}^{t_{iter}}$ to all potential intentions $\mathcal{I}^{t_i}$ searched out at the moment of intention $t_i = t_k - \tau_{his}$ in $\{t_0, \cdots, t_k\}$ in constraints $\mathcal{C}$, where τ_{his} presents a historical period, $t_{iter} \in \{t_i, \cdots, t_k\}$ represents the planning starting moment at which the current iteration calculation takes place. At the initial iteration, we search out all potential destinations from x^{t_i} by accessing to HD map to get road rules and other contextual information in $\mathcal{C}_\mathcal{M}$ as intentions $\mathcal{I}^{t_i}$. Then, according to kinematic and environmental constraints in $\mathcal{C}$, we can generate a reliable trajectory $\mathcal{T}_n^{t_{iter}}$ from $\mathcal{X}^{t_{iter}}$ for each intention $\mathcal{I}_n^{t_i} \in \mathcal{I}^{t_i}$ by a physics-based trajectory generator.

b) Learning-based trajectory matcher: the matcher $M : (\mathcal{X}^{t_{iter}}, \mathcal{T}^{t_{iter}}) \mapsto \bar{\mathcal{P}}^{t_{iter}}$ is employed to evaluate each trajectory $\mathcal{T}_n^{t_{iter}} \in \mathcal{T}^{t_{iter}}$ generated by G to get corresponding probability $\bar{\mathcal{P}}_n^{t_{iter}} \in \bar{\mathcal{P}}^{t_{iter}}$, where $\bar{\mathcal{P}}^{t_{iter}}$ presents the match probabilities produced by a single observation. Notably, this component is only used to match each potential predicted trajectory $\mathcal{T}_n^{t_{iter}} = \{x_n^t|t = t_{iter}, \cdots, t_k, \cdots, t_k + \tau_{pre}\}$ with observed states $\mathcal{X}^{t_{iter}} = \{x^t|t = t_{iter}, \cdots, t_k\}$ to get corresponding match probability $\bar{\mathcal{P}}_n^{t_{iter}}$, rather than regressing future motion information $\{x^t|t = t_{k+1}, \cdots, t_k + \tau_{pre}\}$ as most learning-based methods do.

c) Kalman-based intention estimator: the estimator $E : (\mathcal{P}^{t_{iter-1}}, \bar{\mathcal{P}}^{t_{iter}}) \mapsto \mathcal{P}^{t_{iter}}$ takes charge of obtaining the optimal estimation $\mathcal{P}^{t_{iter}}$ in a temporal sequence by Markov assumption. This component is optional but it is crucial for autonomous driving application, as it can improve the consistency of intentions and the accuracy of trajectories, making system more robust.

III. Method

This section details the prediction process of the IMTP method. The main steps are shown in Algorithm 1.

Algorithm 1: The IMTP method

Inputs : observed states $\mathcal{X}^{t_i} = \{x^t|t = t_i, \cdots, t_k\}$, all constraints $\mathcal{C} = \mathcal{C}_\mathcal{M} \cup \mathcal{C}_\mathcal{V}$

Outputs: multi-modal predicted trajectories with corresponding probabilities $\mathcal{R}^{t_k} = \{(\mathcal{T}_n^{t_k}, \mathcal{P}_n^{t_k})|n = 1, 2, \cdots, N\}$

```
Preprocessing: search out all potential intentions I^{t_i} by (X^{t_i}, C)
for t_iter ← t_i to t_k do
    Trajectory Generating:
        G : (X^{t_iter}, I^{t_i}, C) ↦ T^{t_iter}
    Intention Matching:
        M : (X^{t_iter}, T^{t_iter}) ↦ P̄^{t_iter}
    if t_iter == t_i then
        Initialization: P^{t_iter} ← P̄^{t_iter};
    else
        Kalman Estimating:
            E : (P^{t_iter-1}, P̄^{t_iter}) ↦ P^{t_iter};
    end
end
Results: update R^{t_k} with (T^{t_k}, P^{t_k})
```

A. Preprocessing

As the target agent is uncontrollable and the intention of the target agent can never be entirely predicted, the preprocessing step aims to identify all potential target states, referred to as intentions, from the HD map [15] illustrated in Fig. 2. These intentions encompass target reference paths and target speeds for subsequent planning tasks.

Given an assumption that traffic participants will tend to stay in one lane, and the driving intention will not change in

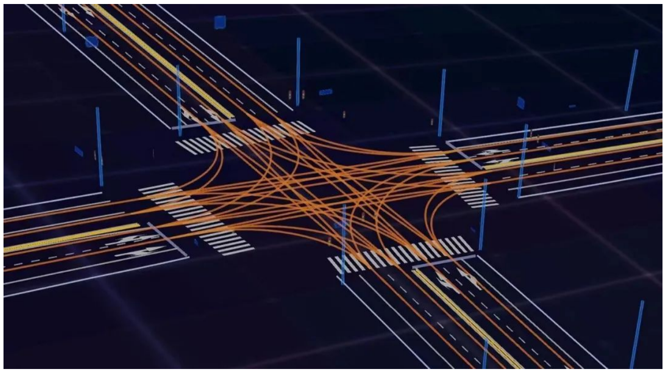


Fig. 2: An Example of A HD Map. This figure shows a semantic vector map with useful lane-level detail, such as lane centerlines, traffic direction, and intersection annotations.

a short period, which is true in most situations, especially on highways, this preprocessing step $(\mathcal{X}^{t_i}, \mathcal{C}) \mapsto \mathcal{I}^{t_i}$ involves a heuristic search for potential target lanes with traffic connectivity to the state $\mathcal{X}^{t_i}$ in the HD map. Furthermore, we can derive target reference paths by utilizing the centerlines and their offset lines associated with these potential target lanes. Finally, an estimation of possible target speeds is made based on the kinematic constraints $\mathcal{C}_{\mathcal{V}}$ and the road speed limit from the HD map $\mathcal{C}_{\mathcal{M}}$. After completing these steps, we obtain all possible target states, constituting the intentions within our proposed method.

B. Trajectory Generating

In this step, we formulate our trajectory generation as an optimization problem [16] and solve it by referring the method mentioned in [17]. After preprocessing, we now have required inputs of $G : (\mathcal{X}^{t_{iter}}, \mathcal{I}^{t_i}, \mathcal{C}) \mapsto \mathcal{T}^{t_{iter}}$ to generate all reachable trajectories for corresponding potential intentions. Unlike classic physics-based methods [4]–[6], for leaving enough observations for intention matching, the trajectory generation does not use the final observed state x^{t_k} as the start state, but the state of the moment before the final observed moment $x^{t_{iter}}|t_{iter} \in \{t_i, \cdots, t_k\}$.

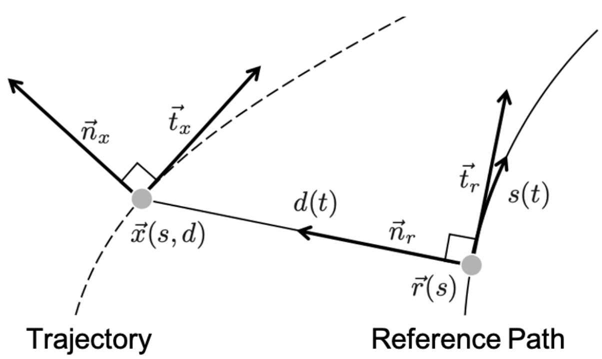


Fig. 3: Trajectory Generation in A Frenet Frame

According to the trajectory generation method mentioned in [17], we can generate a kinematic trajectory from a certain start state to a reference path with a certain end state. In our method, $x^{t_{iter}}$ is the start state and reference paths are in different intentions $\mathcal{I}^{t_i}$. As shown in Fig. 3, the Frenet frame is a dynamic curvilinear frame with a tangential vector $\vec{t}_r$ and a normal vector $\vec{n}_r$ at a certain point $r(s)$ on the lane centerline. The Cartesian coordinate $\vec{x}(x, y)$ could be easily converted to the Frenet coordinate $\vec{x}(s, d)$, with the relation

$$\vec{x}(s(t), d(t)) = \vec{r}(s(t)) + d(t) \cdot \vec{n}_r(s(t)) \tag{1}$$

in which $\vec{r}(s(t))$ represents a dynamic vector pointing from the path root, $s(t)$ and $d(t)$ denote the covered arc length and the perpendicular offset, both on a dynamic frame at time t.

In the prediction problem, we are more concerned about the lateral movement $d(t)$, but are not sensitive to the longitudinal movement $s(t)$. Because lateral movement determines the road selection, longitudinal movement only determines the speed of vehicle. Meanwhile, considering that human perception obviously weights lateral and longitudinal changes of acceleration differently $\dddot{s}(t)$ and we want to fast convergence in a period of T to reference path $d_{t_k+\tau_{pre}} \sim 0$ without collision from t_{iter} to $t_{t_k+\tau_{pre}}$, we get the cost function of lateral movement

$$C_{lat} = w_j \int_{t_{iter}}^{t_{t_k+\tau_{pre}}} \dddot{s}(t)^2 dt + w_t T + w_d {d_{t_k+\tau_{pre}}}^2 \tag{2}$$

with weight coefficients $w_j, w_t, w_d > 0$.

In the longitudinal movement, we ignore some high-order state variables and get the cost function

$$\begin{aligned} C_{lon} = w_j &\int_{t_{iter}}^{t_{t_k+\tau_{pre}}} \dddot{s}(t)^2 dt \\ &+ w_t T + w_v (\dot{s}_{t_k+\tau_{pre}} - \dot{s}_{tar})^2 \end{aligned} \tag{3}$$

with weight coefficients $w_j, w_t, w_v > 0$.

In equation (2) and (3), there are also some constraints including the start state, process constraints and kinematic constraints that need to be satisfied. The start state $[s_{iter}, \dot{s}_{iter}, \ddot{s}_{iter}, d_{iter}, \dot{d}_{iter}, \ddot{d}_{iter}]$ can be easily gained from observation. Due to the inaccuracy of the observation, the higher order can be set zero as $[s_{iter}, \dot{s}_{iter}, 0, d_{iter}, \dot{d}_{iter}, 0]$. For process constraints, we set target velocity $\dot{s}_{tar}$ to the minimum of a guessed speed and road speed limit from HD map and leave s_{tar} unconstrained.

$$\dot{s} \in [0, \dot{s}_{tar}] \tag{4}$$

For kinematic constraints, we have

$$\alpha \in [-\alpha_{max}, \alpha_{max}] \tag{5}$$

$$\kappa \in [0, \kappa_{max}] \tag{6}$$

where the maximum value α_{max} for acceleration α and κ_{max} for curvature κ are control parameters of vehicles set by us.

C. Intention Matching

As we have obtained all trajectories $\mathcal{T}^{t_{iter}}$ to potential intentions $\mathcal{I}^{t_i}$ from a moment t_{iter} in the past, we can confirm the potential intention $\mathcal{I}_n^{t_i}$ by matching each corresponding predicted trajectory from calculation starting moment to the final observed moment $\{x_n^t | t = t_{iter}, \cdots, t_k\} \in \mathcal{T}^{t_{iter}}$

with observed states $\mathcal{X}^{t_{iter}} = \{x^t | t = t_{iter}, \cdots, t_k\}$ to get corresponding match probability $\bar{\mathcal{P}}_n^{t_{iter}}$.

The matching problem is a typical classification problem. In recent years, with the rapid development of machine learning methods, the design of classifiers suitable for different data and targets is changing rapidly. Considering that this paper focuses on the novel framework proposed, we choose the classical classifier Support Vector Machine (SVM) to implement it [18]. In terms of data preparation, we use delta data between the predicted trajectory and the observed trajectory to calculate feature vectors so as to improve the representativeness of features, thus ensuring data consistency. The specific data include delta time, delta heading angle, delta velocity and offset of their position which are all sampled at the same timestamp.

In the training dataset, we use the same method to obtain the data feature vector, but calculate endpoint error with ground truth directly, so as to label "the matched" for those errors lower than a threshold, and others mark as "the unmatched". In the testing dataset, we use the trained model to obtain its matching probability according to the calculated feature vector.

D. Kalman Estimating

Given the predicted probability of intentions and their trajectories $\bar{\mathcal{P}}^{t_{iter}}$ by intention matching at t_{iter}, we can use the Kalman-based estimator $E : (\mathcal{P}^{t_{iter-1}}, \bar{\mathcal{P}}^{t_{iter}}) \mapsto \mathcal{P}^{t_{iter}}$ to obtain optimal probability estimation at t_{iter} by Markov assumption.

In Kalman optimal estimation iterations, we choose the vector $\mathcal{P}^{t_{iter}}$ of probabilities of N predicted trajectories as state. Assume $\mathcal{P}^{t_{iter}}$ follows a normal distribution, then we can get the optimal probability estimate at t_k by recursion from t_i.

$$\mathcal{P}^{t_{iter}} = (1 - K_{gain}^{t_{iter}})\mathcal{P}^{t_{iter-1}} + K_{gain}^{t_{iter}}\bar{\mathcal{P}}^{t_{iter}} \tag{7}$$

where $K_{gain}^{t_{iter}}$ is the Kalman gain in iterations, a value between $[0, 1]$ in our case.

The downstream planning task is able to use its corresponding intention and trajectory according to the probability.

IV. Experiments and Results

In this section, the dataset used in the experiment, the evaluation methods, and the final comparison results will be presented to demonstrate the advantages of the proposed method in accuracy and robustness.

A. Dataset

The Argoverse Motion Forecasting dataset [14] is one of the largest benchmarks for trajectory prediction, including 327,790 sequences of interesting scenarios. Each sequence follows the trajectory of the main agent for 5 seconds while keeping track of all other actors (e.g. car, pedestrian). Each 5-second sequence contains the centred locations of each tracked agent sampled at 10 Hz, in which one vehicle with relatively complex motion is marked as the prediction target. The reason for choosing this dataset is that Argoverse includes HD maps with 290 km of mapped lanes with geometric and semantic metadata which provides us with essential information for identifying intentions. In our experiment, the 5-second trajectory is split into the first 2 seconds for observation and the last 3 seconds for verifying prediction. Considering that the IMTP method needs to achieve optimal estimation through Kalman iteration, we will iterate from the first frame of 2-second observation data to the 20th frame to identify the final trajectory.

B. Evaluation Method

Prediction is a very difficult task, so for the safety of autonomous driving, multiple possible prediction trajectories are more important than the single most likely one. Our method can generate multi-modal future trajectories, so in the evaluation, we use the descending order of the probability of predicted trajectories to sort and evaluate the accuracy of the best K predicted trajectories, where $K = 1, 3, 6$ in Argoverse.

We follow the evaluation criteria in the dataset to perform a comparable analysis of the results. Minimum Average Displacement Error (minADE) and Minimum Final Displacement Error (minFDE) of best $K = 1, 3, 6$ predicted trajectory are used to evaluate how good is the best. The minADE is the average $L2$ distance error of the best predicted trajectory. The minFDE is the $L2$ distance error of the best predicted trajectory at the final timestamp. Notably, the best predicted trajectory refers to the predicted trajectory with the minimum endpoint error. In other words, minADE refers to Average Displacement Error (ADE) of the trajectory which has minimum Final Displacement Error (FDE), and not minimum ADE, since we want to evaluate the single best forecast. Miss Rate (MR) is the ratio of scenarios that none of K predicted trajectories has less than 2 meters $L2$ FDE. So we can see that the minFDE and the minADE are used to evaluate the accuracy of the trajectory prediction, while the MR is used to evaluate the robustness of the trajectory prediction.

C. Results

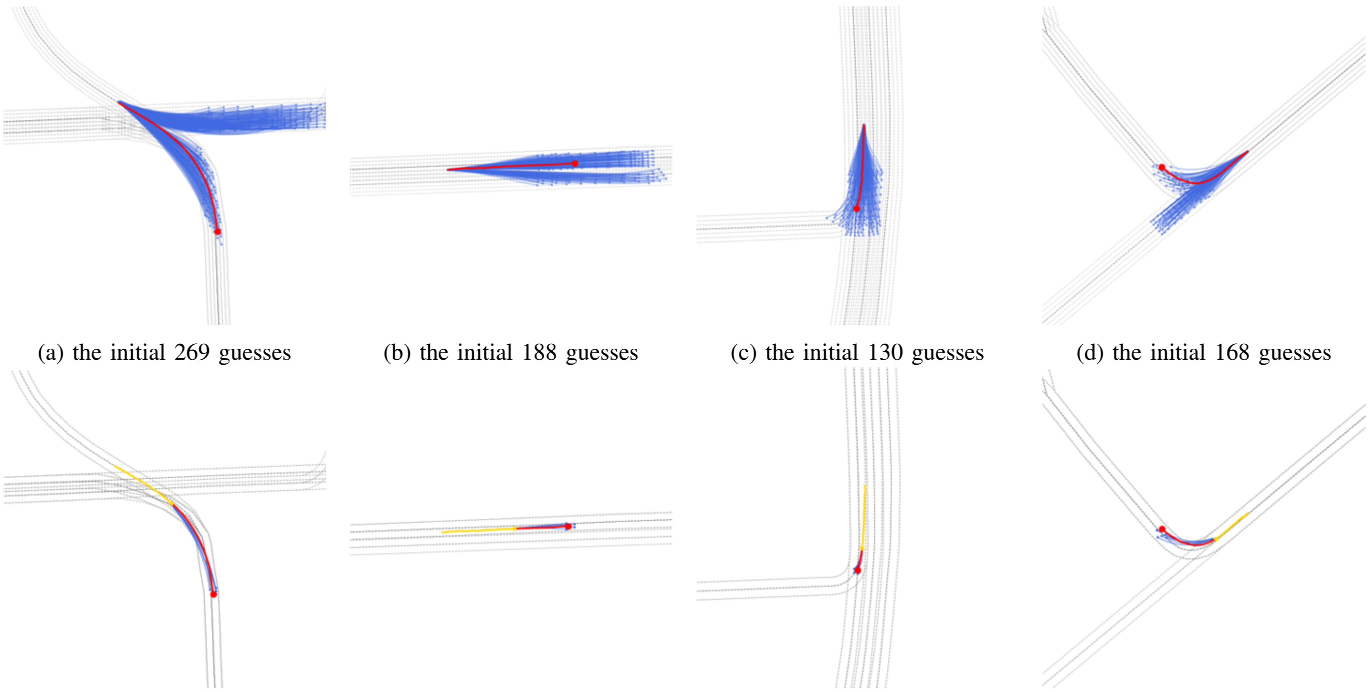
Comparing our IMTP method with the baseline provided by the Argoverse dataset, the TABLE I demonstrates a distinct advantage. In contrast to three methods used in the database baseline—Constant Velocity (CV) [4], Nearest Neighbor (NN) with map [19], and Long Short-term Memory (LSTM) with map [20]—our method consistently produces accurate multi-modal predictions across diverse scenarios. Notably, the MR significantly outperforms the baseline, highlighting the robustness of our approach.

Fig. 4 showcases various trajectory prediction scenarios, such as lane changing and shunting, where our method accurately identifies possible intentions, ultimately converging towards the target direction and providing precisely predicted trajectories.

However, the MR suggests that there is still room for improvement when deploying our framework in a real autonomous driving system, considering that intentions in real scenarios may deviate from our assumptions.

TABLE I: Comparison with the Argoverse Baselines

Methods	K = 1			K = 3			K = 6		
	minADE (m)	minFDE (m)	MR	minADE (m)	minFDE (m)	MR	minADE (m)	minFDE (m)	MR
Argo-CV	3.53	7.89	0.83	-	-	-	-	-	-
Argo-NN+map	3.65	8.12	0.94	3.01	6.43	0.80	2.6	5.32	0.75
Argo-LSTM+map	2.92	6.45	0.75	2.31	4.85	0.71	2.08	4.19	0.67
IMTP (ours)	**1.71**	**2.85**	**0.30**	**1.65**	**2.70**	**0.28**	**1.63**	**2.65**	**0.27**



(a) the initial 269 guesses (b) the initial 188 guesses (c) the initial 130 guesses (d) the initial 168 guesses

(e) the best 6 candidates in (a) (f) the best 6 candidates in (b) (g) the best 6 candidates in (c) (h) the best 6 candidates in (d)

Fig. 4: Results of Typical Scenes. The initial guesses and final predicted trajectories of the same scene are displayed in a top-down manner. The first row shows the search results of all potential intentions. Red represents the real trajectory. Blue represents all possible intentions and their trajectories. The final output result is shown in the second row. The orange trajectory represents the observed 2s. Red represents ground truth for the next 3 seconds and blue represents the multiple predicted trajectories for those 3s. In both the upper and lower rows, the dark grey line represents the lane center line, while the light grey line represents the reference path in the first row and the lane boundary line in the second row.

V. Conclusions and Future Works

We propose a prediction framework, IMTP, which uses a physics-based trajectory generator to generate vehicle trajectories according to potential intentions, and then matches and selects intentions based on existing observations. The IMTP guarantees the feasibility of trajectories by utilizing a physics-based trajectory generator to generate future trajectories under explicit constraints. Meanwhile, a learning-based matcher is used to capture implicit interactions among traffic participants and gives probabilities of each intention. It makes more robust predictions by using a Kalman-based estimator to keep updating the probability and filter out the best future trajectory. The proposed framework effectively integrates the physical vehicle model, road-related environmental factors, and the interactions of the surrounding vehicles. With its novel framework design, the IMTP outperforms existing technologies in terms of prediction accuracy, feasibility and robustness. In addition, this framework is particularly suited for closed road scenarios like highways, benefiting from clear lane markings and a single type of traffic participant, thereby enhancing the system's ability to respond promptly to obstacles, avoid collisions, and improve overall comfort.

While this work presents the prediction of vehicle intentions and trajectories, the specific method heavily relies on HD map to determine intentions. Nevertheless, the core idea of the proposed method can be applied to any agent trajectory prediction where potential intentions can be identified. Furthermore, because this paper focuses on demonstrating the new framework IMTP, there is still a lot of room for enhancing the design of the learning-based matcher, which could significantly improve the prediction effectiveness and broaden its applicability in specific domains. All these ongoing advancements will further refine the concepts presented in this framework and enable better prediction outcomes, ultimately increasing its versatility across various applications.

REFERENCES

[1] C.-Y. Chan, "Advancements, prospects, and impacts of automated driving systems," *International journal of transportation science and technology*, vol. 6, no. 3, pp. 208–216, 2017.

[2] Y. Huang, J. Du, Z. Yang, Z. Zhou, L. Zhang, and H. Chen, "A survey on trajectory-prediction methods for autonomous driving," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 3, pp. 652–674, 2022.

[3] F. Leon and M. Gavrilescu, "A review of tracking and trajectory prediction methods for autonomous driving," *Mathematics*, vol. 9, no. 6, p. 660, 2021.

[4] S. Ammoun and F. Nashashibi, "Real time trajectory prediction for collision risk estimation between vehicles," in *2009 IEEE 5Th international conference on intelligent computer communication and processing*. IEEE, 2009, pp. 417–422.

[5] A. Houenou, P. Bonnifait, V. Cherfaoui, and W. Yao, "Vehicle trajectory prediction based on motion model and maneuver recognition," in *2013 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2013, pp. 4363–4369.

[6] J. Ziegler, P. Bender, M. Schreiber, H. Lategahn, T. Strauss, C. Stiller, T. Dang, U. Franke, N. Appenrodt, C. G. Keller *et al.*, "Making bertha drive—an autonomous journey on a historic route," *IEEE Intelligent transportation systems magazine*, vol. 6, no. 2, pp. 8–20, 2014.

[7] R. Schubert, E. Richter, and G. Wanielik, "Comparison and evaluation of advanced motion models for vehicle tracking," in *2008 11th international conference on information fusion*. IEEE, 2008, pp. 1–6.

[8] N. Lee, W. Choi, P. Vernaza, C. B. Choy, P. H. Torr, and M. Chandraker, "Desire: Distant future prediction in dynamic scenes with interacting agents," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 336–345.

[9] Z. Zhao, H. Fang, Z. Jin, and Q. Qiu, "Gisnet: Graph-based information sharing network for vehicle trajectory prediction," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–7.

[10] J. Ngiam, V. Vasudevan, B. Caine, Z. Zhang, H.-T. L. Chiang, J. Ling, R. Roelofs, A. Bewley, C. Liu, A. Venugopal *et al.*, "Scene transformer: A unified architecture for predicting future trajectories of multiple agents," in *International Conference on Learning Representations*, 2021.

[11] H. Song, D. Luan, W. Ding, M. Y. Wang, and Q. Chen, "Learning to predict vehicle trajectories with model-based planning," in *Conference on Robot Learning*. PMLR, 2022, pp. 1035–1045.

[12] H. Girase, H. Gang, S. Malla, J. Li, A. Kanehara, K. Mangalam, and C. Choi, "Loki: Long term and key intentions for trajectory prediction," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 9803–9812.

[13] K. Mangalam, Y. An, H. Girase, and J. Malik, "From goals, waypoints & paths to long term human trajectory forecasting," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 15 233–15 242.

[14] M.-F. Chang, J. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan *et al.*, "Argoverse: 3d tracking and forecasting with rich maps," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 8748–8757.

[15] A. Barsi, V. Poto, A. Somogyi, T. Lovas, V. Tihanyi, and Z. Szalay, "Supporting autonomous vehicles by creating hd maps," *Production Engineering Archives*, vol. 16, 2017.

[16] A. Gasparetto, P. Boscariol, A. Lanzutti, and R. Vidoni, "Path planning and trajectory planning algorithms: A general overview," *Motion and Operation Planning of Robotic Systems: Background and Practical Approaches*, pp. 3–27, 2015.

[17] M. Werling, J. Ziegler, S. Kammel, and S. Thrun, "Optimal trajectory generation for dynamic street scenarios in a frenet frame," in *2010 IEEE international conference on robotics and automation*. IEEE, 2010, pp. 987–993.

[18] P. Kumar, M. Perrollaz, S. Lefevre, and C. Laugier, "Learning-based approach for online lane change intention prediction," in *2013 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2013, pp. 797–802.

[19] Y. Xing, C. Lv, and D. Cao, "Personalized vehicle trajectory prediction based on joint time-series modeling for connected vehicles," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 1341–1352, 2019.

[20] S. H. Park, B. Kim, C. M. Kang, C. C. Chung, and J. W. Choi, "Sequence-to-sequence prediction of vehicle trajectory via lstm encoder-decoder architecture," in *2018 IEEE intelligent vehicles symposium (IV)*. IEEE, 2018, pp. 1672–1678.