Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Towards Incorporating Pedestrian Intention Predictions into Behavior Planning using Virtual Reality Co-Simulators

Angie Nataly Melo Castillo, Markus Amann, Carlota Salinas Maldonado, Maytheewat Aramrattana, Thomas H Weisswange, Malte Probst , Miguel Ángel Sotelo

2025

Preprint:

This is an accepted article published in 36th IEEE Intelligent Vehicles Symposium (IV 2025): 14th Workshop on Human Factors in Intelligent Vehicles & Supporting Vehicle-Pedestrian Interactions. The final authenticated version is available online at: https://doi.org/10.1109/IV64158.2025.11097738 Copyright 2025 IEEE

Towards Incorporating Pedestrian Intention Predictions into Behavior Planning using Virtual Reality Co-Simulators

Angie Nataly Melo*[1], Markus Amann*[1,2], Carlota Salinas[1], Maytheewat Aramrattana[3], Thomas H. Weisswange[2], Malte Probst[2], Miguel Ángel Sotelo[1]

***Abstract*—Interaction modeling plays a huge role in understanding human behavior in traffic. This is especially relevant when it comes to interactions between vehicles and vulnerable road users such as pedestrians. Thus, pedestrian intention prediction is an ongoing field of research in order to understand the pedestrians' decision making. Most state-of-the-art prediction frameworks are trained on large-scale datasets and evaluated with respect to acknowledged benchmarks. These datasets lack the ability to account for the reciprocal nature of interactions between pedestrians and vehicles and the effects of the two agents influencing each other. In this work, we demonstrate first steps towards assessing pedestrian prediction algorithms within realistic scenarios including the interaction effects arising from its interplay with a planning component. For this, we validate an existing prediction framework trained on benchmark datasets with situations from a virtual reality (VR) pedestrian-vehicle co-simulator that allows us to include the effect of vehicle planning on pedestrian behavior. We evaluate the performance of the prediction framework comparing data from pre-recorded real-world datasets with data from our co-simulation study and conduct an ablation analysis to identify the most important features for pedestrian intention prediction. The results highlight the significance of pedestrian action and proximity to the road.**

***Index Terms*—Pedestrian-Vehicle Interaction, Pedestrian Prediction, Virtual Reality Co-Simulation**

I. Introduction

The increasing complexity of urban environments and the widespread integration of autonomous systems highlight the pressing need for human-centric adaptive interaction frameworks. In scenarios where automated vehicles (AVs) interact with pedestrians, understanding and predicting pedestrian intentions is critical for ensuring safety and optimizing behavior planning. Pedestrian prediction serves a large range of applications, such as inferring the intention of the pedestrian in certain traffic situations enabling safe behavior planning, or creating behavioral models for simulated agents [1]. However, current systems often struggle with the unpredictability of pedestrian actions, particularly in dynamic and densely populated areas.

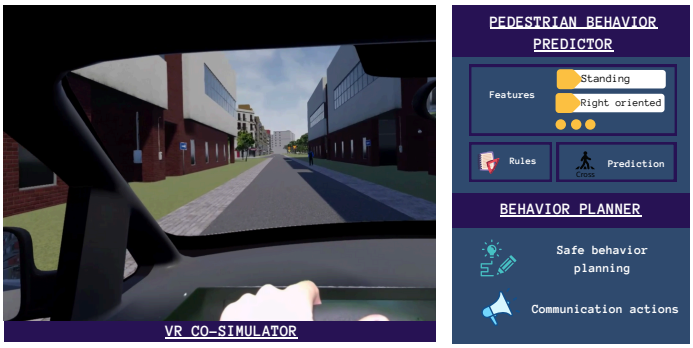

Fig. 1. Individual components to incorporate pedestrian intention prediction in VR co-simulator into behavior planning.

The European project Holistic and adaptivE Interface Design for human-technology Interactions (HEIDI) emphasizes the development of holistic and adaptive interface designs to enhance human-technology interactions, aiming to create systems that are intuitive, reliable, and context-aware. Building on this vision, our research addresses a pivotal challenge: integrating pedestrian intention predictions into autonomous decision-making processes. By leveraging VR simulators, we can create controlled, immersive environments that mimic real-world conditions, enabling the collection of high-quality behavioral data and the testing of predictive models while ensuring safe conditions. However, the suitability of prediction algorithms that have been developed using benchmark datasets for the application in VR simulators requires to be validated.

In this work, we employ an existing pedestrian intention prediction framework [2] to interaction data of vehicle-pedestrian interactions from a VR co-simulator. The algorithm employs a neuro-symbolic approach that combines deep learning with fuzzy logic, utilizing a set of explainable features. We demonstrate the adaptability of the algorithm, which was trained on benchmark datasets (Joint Attention for Autonomous Driving (JAAD) and Pedestrian Intention Estimation (PIE)), to the virtual settings in the co-simulator. Moreover, we conduct a feature ablation analysis to investigate the relevance of the

This research was supported by the HEIDI project which has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101069538. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

*The authors contributed equally to this work.

[1]Computer Engineering Department, University of Alcalá, 28805 Alcalá de Hernares, Madrid, Spain, Email:{nataly.melo, carlota.salinasmaldo and miguel.sotelo}@uah.es

[2]Honda Research Institute Europe GmbH, Carl-Legien-Str. 30, 63073 Offenbach, Germany, Email:{firstname.lastname}@honda-ri.de

[3] Swedish National Road and Transport Research Institute (VTI), 583 30 Linköping, Sweden, Email: maytheewat.aramrattana@vti.se

individual features comparing the results from the benchmark dataset with the results from the co-simulator.

This approach not only aligns with HEIDI's goals of fostering adaptive and interactive systems but also introduces a novel methodology for advancing safety and cooperation in human-autonomous interactions. By bridging predictive modeling with behavior planning, our work seeks to contribute to the development of technology that seamlessly integrates into human-centered environments while prioritizing trust, adaptability, and efficiency. Figure 1 depicts the main components of this study around the HEIDI context.

The rest of the article is organized as follows: Section II discusses related work regarding pedestrian intention prediction. Details about the co-simulation setup for data collection, the standardized data interfaces and the pedestrian intention predictor are introduced in Section III. Subsequently, the results are described and discussed in Section IV. In section V, we discuss the applicability of prediction frameworks, as presented in this work, as model extension for interactive behavior planning. Finally, Section VI concludes the work and provides insights for future research.

II. Related Work

In traffic interactions, the agents' behavior can be categorized with respect to their motivation for the outcome of the interaction [3]. Mutualistic and altruistic behavior improve the costs of the interaction partner, whereas selfish and spiteful behavior deteriorate the outcome for the interaction partner. However, understanding the intentions of other agents, thus, being able to successfully resolve interactions, requires reliable and accurate predictions. Pedestrian intention prediction, also referred to as crossing action prediction, is a critical machine learning (ML) task aiming at forecasting whether a pedestrian will cross a road at a future time. This task is pivotal for enhancing the safety and functionality of autonomous vehicles and intelligent transportation systems. Researchers have explored a diverse array of algorithms and architectures to address this challenge. In [4], the authors argue that pedestrian trajectory predictions based on kinematic features ignore important contextual information. However, contextual models which are designed for specific scenarios, such as interactions at pedestrian crosswalks, tend to over predict crossings events and are not suitable for generalized scenarios. Among the notable approaches is SingleRNN, which employs recurrent neural networks (RNNs) to model sequential data and capture temporal dependencies in pedestrian behaviors [5]. Similarly, CapFormer, leveraging the transformer architecture with a self-attention mechanism, excels at capturing complex spatiotemporal patterns [6]. More sophisticated frameworks like Stacked RNN with Multilevel Fusion [7] and Convolutional LSTM [8] incorporate multi-layered learning and spatial-temporal feature integration to enhance predictive accuracy.

While substantial progress has been made in developing models for pedestrian crossing prediction, relatively few studies address the issue of explainability. Explainability in ML is crucial for understanding the factors driving predictions, particularly in safety-critical domains like autonomous driving. For instance, transformers, as highlighted in [9], provide interpretability advantages through their attention mechanisms, which allow insights into the feature importance of predictions.

III. Methodology

This section provides detailed information on data collection from the co-simulation environment, the data interfaces utilized in this work and the methodology for the pedestrian intention predictor.

A. Driver-Pedestrian Co-Simulation Setup for Data Collection

Training and testing prediction algorithms on large datasets with respect to well-known and acknowledged benchmarks is a crucial step for creating a reliable intention estimation. However, it is equally important to further validate its applicability in real world situations, especially since replaying pre-recorded data does not account for the closed-loop characteristics of traffic interactions when it comes to incorporating the prediction into the behavior planning step. Despite the behavioral gap between interactions in VR and real-world, VR simulators offer a suitable method for simulating realistic interactions between human road users in a safe and controlled environment [1], [10].

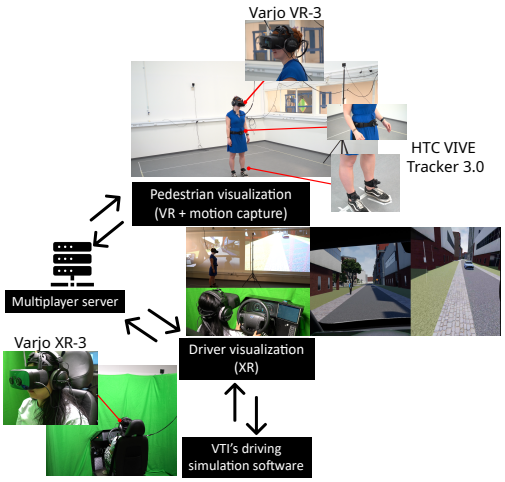

Fig. 2. Overview of the simulators used for data collection

The pedestrian intention prediction algorithm is applied to data that was collected in a VR co-simulation framework [11], where pedestrians can safely interact with human drivers. Figure 2 shows an overview of the co-simulation framework at the Swedish National Road and Transport Research Institute (VTI), which was used for data collection. In this simulation framework, the driver and pedestrian are put in the same simulation environment, thus, can interact with each other in real-time. The pedestrian is equipped with a Varjo VR-3 headset and three HTC VIVE 3.0 trackers (two at ankle joints, one around waist). The position and orientation of the pedestrian are then tracked using the headset together with the trackers, which allow us to represent the motion of the pedestrian to the driver through an animated avatar. Similarly,

the driver is equipped with a Varjo XR-3 headset, which allows the driver to see their own hands, the actual steering wheel and dashboard while the rest of the environment is simulated. Likewise, the position and orientation of the car are tracked and visualized to the pedestrian. Hence, this simulation setup allows realistic interactions between human drivers and pedestrians in a simulated VR environment.

Data used in this paper is collected from 35 pairs of drivers and pedestrians using the co-simulation framework presented above. In total, we collect data from 560 interactions in the experiment, which are used to validate the pedestrian intention predictor framework. During the data collection, the drivers drive the car through an urban traffic environment entering a single lane road in which several pedestrians are present. They can freely decide how to drive and whether to yield to the pedestrians. One of these pedestrians is the avatar of the interacting pedestrian which is controlled by another participant (equipped with the VR-headset and the trackers). The pedestrian participants have the task to reach the other side of the road upon receiving a certain signal. However, after the signal is triggered, they are free to decide whether they want to cross before the vehicle or after.

B. Standardized Data Interfaces

The pedestrian position and orientation data, as well as the kinematic vehicle data and the vehicle position are collected by a multiplayer server which processes the data according to the definition of standardized interfaces. The multiplayer server then broadcasts the data to all other client modules. The interfaces comprise three different main classes. The first class contains information about the vehicle position in the virtual environment. This class supports both local or global map data. The second class contains all data about the vehicle kinematics, such as velocity, acceleration, and yaw rate. The third class provides all input about surrounding traffic participants with respect to the ego vehicle. The standardized interfaces are based on the .msg format for ROS messages. This setup enables a high level of modularity since various modules, such as the co-simulator, the vehicle behavior planner, and the pedestrian prediction module share the same interfaces and use the same set of messages. Additionally, this ROS compatible convention supports automatic code generation for several programming languages (Python, C++, ...). This facilitates connecting modules written in different programming languages. Moreover, using standardized interfaces allows for switching the data input to the prediction framework between different data sources, such as online data input in real world, online data input in the simulator or recorded data.

C. Pedestrian Intention Prediction Framework

In the context of interactions between pedestrians and drivers, or automated driving systems, understanding and predicting pedestrian behavior is a critical factor. Grasping the rationale behind specific decisions can significantly enhance road safety and reduce traffic accidents. Moreover, explaining ego-vehicle decisions provides valuable descriptive information about the causal history of actions taken. To address these needs, we developed prediction methods based on knowledge representation to anticipate pedestrians' most likely future actions in a given contextual situation.

The method that is used in this work employs a neuro-symbolic model that integrates deep learning and fuzzy logic. As illustrated in Figure 3, the approach consists of three main steps. First, we extract relevant data from pedestrian action datasets such as JAAD and PIE. This data is then used with the IVTURS-FARC algorithm [12] to generate a set of fuzzy rules encapsulating the knowledge provided by the datasets. Using these rules, we define a Takagi-Sugeno (TS) fuzzy inference system, which enables the representation of non-linear systems through a set of fuzzy rules with linear state equations as consequents. Finally, pedestrian features are provided as input to the TS fuzzy inference system, consolidating the predictor model. The full explanation and further details of the pedestrian intention prediction used in this work can be found in [2].

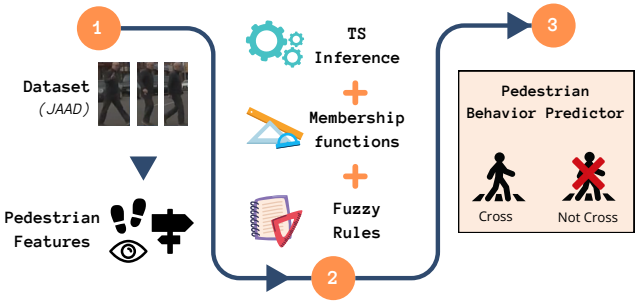

Fig. 3. Pedestrian intention predictor pipeline

For the HEIDI context, six carefully selected features are used as inputs to ensure accurate and reliable predictions:

- **Body orientation.** Represents the pedestrian's posture through an angle ranging from 0° to 360°. *Values:* Left orientation, Right orientation, Align with the car direction, Opposite to the car direction.
- **Action.** Indicates the pedestrian's motion state. *Values:* Walking or Standing.
- **Attention:** Reflects the pedestrian's attention, specifying whether the pedestrian is looking at the ego-vehicle. *Values:* Looking or Not Looking.
- **Proximity to the road:** Indicates whether the pedestrian is near the road. *Values:* Near, Moderate distance, Far.
- **Distance:** Represents the estimated distance between the pedestrian and the ego-vehicle. *Values:* Too near, Near, Moderate distance, Far, Too far.

To develop the neuro-symbolic predictor, we used the JAAD dataset to extract the specified pedestrian features. The Attention and Action features are directly obtained from the dataset annotations, while the remaining features are derived using supplementary neural networks, as outlined in [2].

The integration of the pedestrian intention prediction system within the HEIDI framework involves a comprehensive pro-

cess of data conversion to adapt the information generated in the driver-pedestrian co-simulator. Table I highlights the specific elements extracted from the provided data and utilized as inputs for our neuro-symbolic predictor, ensuring compatibility and effective operation.

TABLE I
SIMULATION STUDY INPUTS CONVERSION

Co-Simulation data	**Unit**	**Pedestrian Feature**	**Data Conversion**
Heading	deg	Body Orientation	Is calculated using the VR's heading
Position x	m	Distance	Is calculated using the Euclidean distance between the VR's position and the ego-vehicle's position
Gaze	deg	Attention	Is computed using the VR's head azimuth angle WRT to the ego-vehicle
Position y	m	Proximity to the road	Is calculated using the VR's position WRT to the road
Velocity x - Velocity y	m/s	Action	Is computed by filtering the VR's velocity

To evaluate the predictor's performance in the co-simulation study, we designed an experimental setup utilizing Precision, Recall, Accuracy, and F1 score metrics.

IV. RESULTS AND DISCUSSION

A. Pedestrian Intention Prediction Results

The model used includes 51 fuzzy rules which help to predict whether a pedestrian would cross or not. To evaluate the performance and anticipation capability of the pedestrian intention predictor, predictions are made to determine whether a pedestrian would cross or not in the 15th, 30th, and 60th subsequent frames. The frame rate is 0.02 seconds.

The results obtained are summarized in Table II. These findings demonstrate that the model achieves its best performance for mid-term predictions (30th frame), reflected in the highest F1 score and a well-balanced trade-off between Precision and Recall. Early predictions (15th frame) exhibit strong Recall, making this timeframe particularly valuable for applications where capturing all potential crossing actions is essential. In contrast, long-term predictions (60th frame) prioritize Precision, effectively minimizing false positives but at the expense of reduced Recall. This trade-off underscores the increasing challenge of accurately forecasting events over extended horizons.

TABLE II
PEDESTRIAN INTENTION PREDICTION RESULTS

Prediction frame	**F1**	**Precision**	**Recall**	**Accuracy**
15th frame	0.72	0.68	**0.76**	**0.91**
30th frame	**0.73**	0.74	0.72	**0.91**
60th frame	0.68	**0.77**	0.61	0.88

Finally, the consistently high Accuracy observed across all prediction frames underscores the model's overall reliability. However, the varying balance between Precision and Recall across different frames highlights the importance of selecting an appropriate prediction horizon based on specific application requirements.

It is important to highlight that, although the results are not statistically significant in terms of numbers, they offer an interesting perspective on how training in real-world environments can be utilized for testing in simulated settings. This suggests that, despite the apparent differences between these environments, the behaviors of pedestrians and drivers can still be effectively represented. Furthermore, it emphasizes the relevance of the conditions and setup of simulation studies when aiming to achieve realistic results and accurately represent road users' behavior in a way that closely mirrors real-world scenarios.

Additionally, we conducted an analysis of the most frequently triggered rules within the study data, focusing on the features that compose each rule. Figure 4 illustrates the activation frequency of each rule, represented as an activation counter that tracks how often a rule was triggered across all frames of the study data. The results reveal that the most prominent features among the top five activated rules are **Proximity** and **Action**, highlighting their significance in pedestrian intention prediction. Notably, there is a clear indication that when the pedestrian's proximity to the road is minimal, there is a high probability of the pedestrian crossing the street. Conversely, when the pedestrian's action is stationary (e.g., standing), it is likely that the pedestrian will not cross. These findings highlight the significant influence of proximity and action in determining pedestrian behavior, underscoring their importance in predictive modeling.

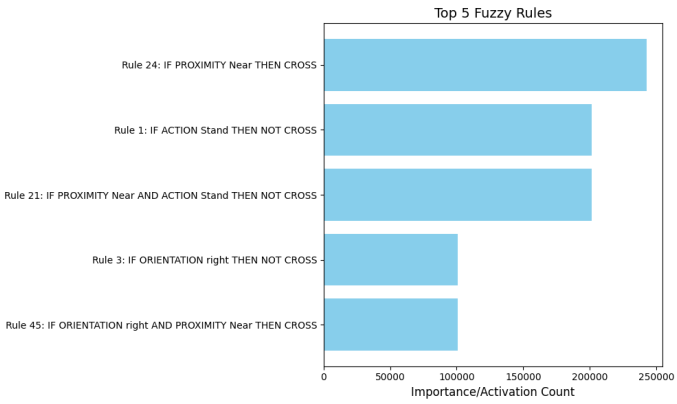

Fig. 4. Most frequently triggered rules within the co-simulation data across the frames

B. Features Ablation Analysis

One of the aims of this study is to identify the most representative features for pedestrian intention prediction and to understand the relationships between them. To achieve this, we conduct a feature ablation study where each of the five features is excluded in turn. For each configuration, we train the model, generate the corresponding fuzzy rules, and test the

model on co-simulated data to predict the pedestrian behavior in the next 30 frames.

Table III summarizes the experiments conducted, highlighting the combination of features used and their corresponding F1, Precision, Recall, and Accuracy metrics. Experiment 1 serves as the baseline, utilizing all five features. This experiment achieves the highest overall performance across all metrics, underscoring the complementary nature of the selected features in predicting pedestrian behavior effectively.

The subsequent experiments reveal that **Proximity** and **Action** are the most critical features, as their absence resulted in the most significant declines in performance metrics. Meanwhile, **Distance** and **Attention** play supportive yet essential roles, contributing to the overall accuracy and robustness of predictions. **Orientation**, while having the least pronounced individual impact, still contributed to overall performance.

However, the experiments demonstrate that no single feature is solely dominant. Instead, pedestrian intention prediction is inherently a multifaceted problem that benefits from the integration of diverse features. The interplay of all five features consistently produced the best results, highlighting the need for a holistic feature set to achieve optimal performance.

V. Discussion of Pedestrian Prediction Modeling for Interactive Behavior Planning

One of the major difficulties in interactive behavior planning is taking into account that prediction and planning are influencing each other [13]. One way to tackle this challenge is by considering multiple behavior options for the interaction partner and planning the own behavior for each of these options [14], [15]. These predictions need to account for semantically different behavior options which reflect the different intentions that the interaction partner could have. This approach allows for breaking the closed loop, which comprises prediction and behavior planning, yet still covering the multi-modality of possible future resolutions. In [14], the authors propose a framework that models the interaction as a negotiation for the passing order through a shared traffic space. In the prediction step, they generate two behavior candidates reflecting the intentions of passing first or passing second through the shared traffic space. Subsequently, they plan the vehicle's behavior for both options with respect to risk, utility and comfort using a probabilistic risk model [16]. This allows planning the vehicle behavior for different possible outcomes, and it enables considering joint costs facilitating cooperative behavior. The approach can be interpreted such that the vehicle repeatedly plans its behavior according to the best joint resolution assuming that the interaction partner will plan their behavior with respect to the same outcome. However, the approach neglects the actual probability if the interaction partner will adhere to the predicted behavior. Moreover, it does not consider the influence of the vehicle behavior on the other agent's future behavior.

Coordinating the agents' actions requires the interaction partners to communicate with each other, which can be achieved in various ways. Implicit communication describes

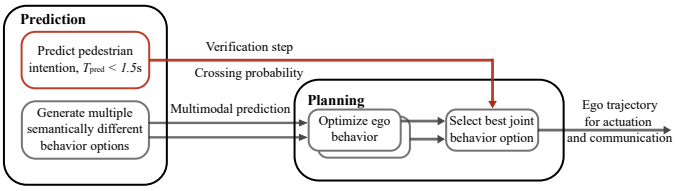

Fig. 5. Extended prediction and planning approach based on [14], [15]. The architecture comprises a verification step whether the pedestrian shows the expected reaction according to the planned behavior and communication.

the information that is implicitly transmitted through the agents' behavior [1], [17]–[19]. Explicit communication can be achieved by additional explicit signals, which can comprise a variety of visual cues, auditory signals, or haptic cues [15], [20]–[23]. Such communication modalities, whether implicit or explicit, change the planning and the action taking of the own behavior. Moreover, communication influences the behavior of the interaction partner, thereby changing the prediction of the future behavior. In vehicle-pedestrian interactions, pedestrian behavior can be characterized as a high-level binary decision of crossing or not crossing before the vehicle. Incorporating a pedestrian intention prediction framework, as presented in this work, into the behavior planner not only facilitates safe behavior planning in reaction to the pedestrian's intention. We hypothesize that the prediction will also portray the effect of the communication on the interaction partner's behavior and crossing decision. Hence, if the vehicle takes communication actions in order to induce a specific pedestrian behavior, we suggest that the prediction can be used to confirm whether the pedestrian reacts as expected. In accordance with [19], the prediction can be interpreted as a verification step. This relation is shown in Figure 5.

VI. Conclusions

This work demonstrated the applicability of a benchmark-trained behavior prediction framework on a co-simulation dataset. The results provide valuable insights, highlighting the potential for training models in real-world environments and applying them effectively in simulated settings. These findings suggest that, despite inherent differences between these environments, pedestrian behavior can be accurately represented and analyzed. Additionally, achieving 91% accuracy in predictions for a minimum of 15 frames in advance underscores the framework's strong performance. Consequently, the robust performance in co-simulation experiments highlights its applicability in both real-world and virtual reality simulators, showcasing its versatility across diverse scenarios.

Moreover, the extraction of diverse pedestrian features from the simulated data plays a critical role in enhancing prediction capabilities. The ablation study underscores the importance of features such as Proximity and Action while also revealing that leveraging a combination of multiple features leads to more effective and accurate pedestrian intention predictions. This multi-feature approach reinforces the need for a holistic perspective when designing prediction frameworks.

TABLE III
PEDESTRIAN INTENTION PREDICTION ABLATION STUDY

Experiment	Rules	Attention	Action	Proximity	Distance	Orientation	F1	Precision	Recall	Accuracy
01	51	X	X	X	X	X	**0.73**	**0.74**	**0.72**	**0.91**
02	36	X	X	-	X	X	0.51	0.76	0.39	0.87
03	22	X	-	X	X	X	0.28	0.16	0.91	0.19
04	38	X	X	X	-	X	0.67	0.69	0.65	0.89
05	43	-	X	X	X	X	0.67	0.72	0.63	0.89
06	28	X	X	X	X	-	0.70	0.68	0.72	0.89

In this work, we describe an approach how pedestrian prediction can be used to improve behavior planning and advance interactions between vehicles and pedestrians. We hypothesize that the prediction can be used to verify that certain vehicle actions induce a specific pedestrian behavior, thus facilitating anticipatory behavior planning.

VII. OUTLOOK

In future work, we will implement the suggested combined prediction and planning approach comprising the behavior verification step. We plan to develop a system that is particularly able to take actions to communicate its intention and future behavior to the pedestrian and account for these actions in the future evolution of the interaction. We will consider the results from this study for the evaluation of the suggested system. Similarly, in future work, we aim to explore knowledge-based approaches for the pedestrian intention predictor and incorporate additional features into the proposed framework.

ACKNOWLEDGMENT

This research has been funded by the HEIDI project of the EC (Grant Agreement 101069538) and the SEGVAUTO5G-CM project of the CAM (reference 2024/ECO-277).

REFERENCES

[1] J. Pekkanen, O. T. Giles, Y. M. Lee, R. Madigan, T. Daimon, N. Merat, *et al.*, "Variable-drift diffusion models of pedestrian road-crossing decisions," *Computational Brain & Behavior*, pp. 1–21, 2022.

[2] A. N. Melo, C. Salinas, and M. A. Sotelo, "Experimental insights towards explainable and interpretable pedestrian crossing prediction," *arXiv preprint arXiv:2312.02872*, 2023.

[3] J. C. Mertens, C. Knies, F. Diermeyer, S. Escherle, and S. Kraus, "The need for cooperative automated driving," *Electronics*, vol. 9, no. 5, p. 754, 2020.

[4] S. Bonnin, T. H. Weisswange, F. Kummert, and J. Schmüdderich, "Pedestrian crossing prediction using multiple context-based models," in *17th International IEEE conference on intelligent transportation systems (ITSC)*. IEEE, 2014, pp. 378–385.

[5] I. Kotseruba, A. Rasouli, and J. K. Tsotsos, "Do they want to cross? understanding pedestrian intention for behavior prediction," in *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 1688–1693.

[6] Lorenzo, Javier and Alonso, Ignacio Parra and Izquierdo, Rubén and Ballardini, Augusto Luis and Saz, Álvaro Hernández and Llorca, David Fernández and Sotelo, Miguel Ángel, "Capformer: Pedestrian crossing action prediction using transformer," *Sensors*, vol. 21, no. 17, 2021.

[7] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Pedestrian action anticipation using contextual feature fusion in stacked rnns," *arXiv preprint arXiv:2005.06582*, 2020.

[8] X. SHI, Z. Chen, H. Wang, D.-Y. Yeung, W.-k. Wong, and W.-c. WOO, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," in *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds., vol. 28. Curran Associates, Inc., 2015.

[9] L. Achaji, J. Moreau, F. Aioun, and F. Charpillet, "Analysis over vision-based models for pedestrian action anticipation," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2023, pp. 5846–5851.

[10] S. Martín Serrano, R. Izquierdo, I. García Daza, M. Á. Sotelo, and D. Fernández-Llorca, "Behavioural gap assessment of human-vehicle interaction in real and virtual reality-based scenarios in autonomous driving," *International Journal of Human–Computer Interaction*, pp. 1–14, 2024.

[11] M. Aramrattana and L. Ochel, "Simulation architecture for real-time driver-pedestrian co-simulation using motion capture," in *Product Solutions Book of the Driving Simulation Conference 2024 Europe VR*. Driving Simulation Association, 2024.

[12] J. Sanz, A. Fernandez, H. Bustince, and F. Herrera, "Ivturs: a linguistic fuzzy rule-based classification system based on a new interval-valued fuzzy reasoning method with tuning and rule selection," *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 3, pp. 399–411, 2013.

[13] S. Hagedorn, M. Hallgarten, M. Stoll, and A. P. Condurache, "The integration of prediction and planning in deep learning automated driving systems: A review," *IEEE Transactions on Intelligent Vehicles*, pp. 1–17, 2024.

[14] R. Wenzel, M. Probst, T. Puphal, T. H. Weisswange, and J. Eggert, "Asymmetry-based behavior planning for cooperation at shared traffic spaces," in *2021 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2021, pp. 1008–1015.

[15] M. Amann, M. Probst, R. Wenzel, and T. H. Weisswange, "Enabling cooperative pedestrian-vehicle interactions using an ehmi," in *2024 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2024, pp. 2962–2969.

[16] M. Probst, R. Wenzel, T. Puphal, M. Komuro, T. H. Weisswange, N. Steinhardt, *et al.*, "Automated driving in complex real-world scenarios using a scalable risk-based behavior generation framework," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2021, pp. 629–636.

[17] A. Dietrich, P. Maruhn, L. Schwarze, and K. Bengler, *Implicit Communication of Automated Vehicles in Urban Scenarios: Effects of Pitch and Deceleration on Pedestrian Crossing Behavior*, 01 2020, pp. 176–181.

[18] K. Tian, A. Tzigieras, C. Wei, Y. M. Lee, C. Holmes, M. Leonetti, *et al.*, "Deceleration parameters as implicit communication signals for pedestrians' crossing decisions and estimations of automated vehicle behaviour," *Accident Analysis & Prevention*, vol. 190, p. 107173, 2023.

[19] M. Amann, M. Probst, R. Wenzel, T. H. Weisswange, and M. Ángel Sotelo, "Optimal behavior planning for implicit communication using a probabilistic vehicle-pedestrian interaction model," *arXiv preprint arXiv:2504.15098*, 2025.

[20] D. Dey, A. Habibovic, A. Löcken, P. Wintersberger, B. Pfleging, A. Riener, *et al.*, "Taming the ehmi jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces," *Transportation Research Interdisciplinary Perspectives*, vol. 7, p. 100174, 2020.

[21] A. Schieben, M. Wilbrink, C. Kettwich, R. Madigan, T. Louw, and N. Merat, "Designing the interaction of automated vehicles with other traffic participants: design considerations based on human needs and expectations," *Cognition, Technology & Work*, vol. 21, pp. 69–85, 02 2019.

[22] A. Habibovic, V. Fabricius, J. Andersson, M. Klingegård, T. Lagström, A. Sirkka, *et al.*, "Communicating intent of automated vehicles to pedestrians," *Frontiers in Psychology*, vol. 9, p. 1336, 08 2018.

[23] W. M. Alvarez, F. M. Moreno, O. Sipele, N. Smirnov, and C. Olaverri-Monreal, "Autonomous driving: Framework for pedestrian intention estimation in a real world scenario," in *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2020, pp. 39–44.