Explainable LSTM-Based Cyclist Intention Prediction at Intersections

Marie-Christin H. Oczko, Osman Tugay Başaran, and Falko Dressler
School of Electrical Engineering and Computer Science, TU Berlin, Germany
{oczko, basaran, dressler}@ccs-labs.org

***Abstract*—Increasing the safety of vulnerable road users (VRUs) in traffic has become a topic of general interest. Predicting cyclists' turning intention in intersections can benefit safety applications in forecasting potential accidents. In this paper, we propose a bidirectional, stacked LSTM intention prediction model utilizing real-world smartphone cycling traces. We show that even imprecise GPS data are sufficient to predict right turns, and straight-going traces with a certainty of 90 % 45 m, and left turns 28 m before the intersection center, resulting in recognizing even the intention of the fastest cyclist in the data set 4.19 s before reaching the center. We further conduct an explainability analysis, including feature engineering, and SHapley Additive exPlanations (SHAP), highlighting the influence of GPS positions, and rotation vectors on our model. Lastly, we investigate the generalizability of our model on untrained intersections, showing first promising results for left turns of 90 % prediction probability 45 m before the intersection center, and probabilities of 90 % 20 m for straight-going traces, for an exemplary intersection.**

***Index Terms*—VRU safety, intention prediction, LSTM networks, explainability, bicycle trajectory prediction**

I. Introduction

In recent years, advanced driver assistance systems (ADASs), autonomous driving, and traffic optimization grew into the role of a game changer for smart cities, intelligent vehicles, and politics. Therefore, ensuring the safety of vulnerable road users (VRUs), especially cyclists, has become a topic of general interest. To achieve this, ADASs have been researched, developed, and deployed in modern cars. However, current solutions do not incorporate VRUs sufficiently [1]. We focus on the safety of such VRUs. Adequate safety systems in both cars and bicycles require accurate trajectory prediction, particularly at intersections, to forecast potential accident scenarios.

Three categories of trajectory prediction can be distinguished [2]: intention prediction, unimodal trajectory prediction, and multimodal trajectory prediction. In the first category, only the general action is predicted, e.g., turning left, or right. The last two categories forecast concrete positions. Unimodal approaches focus on the most likely trajectory, while multimodal compute different paths with their according probabilities. Depending on the objective, a physics- or maneuver-based prediction should be considered. According to Fu et al. [3], physics-based prediction is built on vehicle kinematics models. Here, human factors are either neglected, or the prediction uncertainty is considered by applying the Kalman filter for state estimation. In maneuver-based prediction, driver intention, and possibly environment information are included in the prediction process. In the literature, many approaches combine intention, and uni- or multimodal predictions by using the first class as an input for the other classes of trajectory forecast [4]–[6].

For identifying a driver's intention, various AI techniques can be applied, such as graph neural networks (GNNs), convolutional neural networks (CNNs), or recurrent Neural Networks (RNNs). Some prediction approaches base their predictions on graph neural networks [7]–[9]. Li et al. [8] combine a graph network predicting the social connections with a long short-term memory (LSTM)-based encoder-decoder setup for trajectory prediction. Furthermore, Lee et al. [10] implement a CNN-based solution foretelling lane changes in the context of automatic cruise control. Other works, such as Ding et al. [11], utilize RNNs to compute the probabilistic likelihood of multiple behaviour classes of cars. In the context of RNNs, LSTMs are a common approach when dealing with long time series data, and trajectory predictions [4]–[6], [9]. Using LSTM networks permits to avoid the vanishing gradient problem of RNNs. Furthermore, bidirectional LSTM systems allow capturing context information in both directions [12]. Hence, bidirectional LSTM networks can learn high-level abstractions of sequential features. Additionally, stacked LSTM models can achieve higher prediction accuracy [12]–[14]. Most of these works either focus on cars instead of bicycles, or rely on accurate video data, use sophisticated inertial measurement unit (IMU) setups, and do not research the generalizability of their approaches on untrained intersections.

In this paper, we present a bicycle intention prediction approach utilizing real-world smartphone data from the open-source SimRa toolkit[1] [15]. Similar to the work by Saleh et al. [12], we implement a bidirectional, stacked LSTM network with a sliding window approach to handle trace data of uneven length. However, instead of utilizing video data traces as position input, we use IMU and global positioning system (GPS) data from smartphones, and adapt their design to predict intentions, and conduct an extensive explainability analysis, including feature engineering and SHapley Additive exPlanations (SHAP).[2] Finally, we present a first insight into the generalizability of our approach to untrained intersections.

We show that even imprecise smartphone sensor data are sufficient to predict cyclist intentions. Hence, our solution can be realized using regular smartphones with no additional equipment. By including features such as speed, rotation

[1]https://github.com/simra-project
[2]https://shap.readthedocs.io/en/latest/

vector, and linear acceleration, we show that we can correctly predict intentions to turn left, right, or straight on a four-street intersection with a probability of more than 90 % as early as 28 m before the center of the intersection. We estimate how early we can predict each intention with a 90 % certainty. We demonstrate that even for fast bicycles the time available for sending, and processing a warning message is sufficient. Finally, we present first insights into the generalizability of our approach on untrained intersections, showing we can still achieve a prediction accuracy of 90 % as early as 45 m for left turns, and 20 m before the intersection for going straight.

Our key contributions can be summarized as follows:

- We developed a LSTM-based intention prediction for bicycles utilizing real-world smartphone data.
- We conducted an extensive evaluation of our model, including feature engineering and SHAP.
- We offer first insights on generalizability of our approach by applying it to an untrained intersection.

II. Related Work

Compared to cars, research on the prediction of bicycle trajectories is much more limited. The simplest models are linear computations based on static acceleration, velocity, and heading. For example, Ruß and Naumann [16] extrapolate trajectories to forecast when an object will reach the intersection. More complex approaches, based on data collected via simulation, or IMU sensor data, utilize classifiers or machine learning solutions [17]–[20]. Mathuseck et al. [17] install multiple IMU sensors on a bicycle, and successfully predict intentions such as pedaling, coasting, and braking utilizing a transformer-based classifier. Han et al. [18] forecast maneuvers based on head movements of a cyclist. Zernetsch et al. [19] compare a physical model of a cyclist, as well as a polynomial least-squares approximation in combination with a multilayer perceptron artificial neural network, to the performance of a Kalman Filter. Predicting intentions such as starting, stopping, and passing, their physical model performs 27% better than the Kalman Filter, and the machine learning approach even achieves an improvement of 34% for starting, and stopping. Losada et al. [20] show how the cyclist's behavior impacts the collision probability using bicycle simulator data. They implement simple classifiers to predict whether the cyclist's behavior, e.g., braking, dodging, or accelerating, would lead to or prevent an accident.

Other works utilize camera-based solutions to predict the trajectory of vehicles [9], [12], [21]. For example, Saleh et al. [12] implement a vision-based trajectory prediction by training a bidirectional RNN based on LSTMs. Li et al. [9] include social interactions in their predictions by implementing a spatial-temporal multi-graph network trained for crowded city scenarios, resulting in 9% less prediction error than RNN models in similar setups.

Multiple works, especially in the context of cars, have successfully applied LSTM networks to trajectory prediction approaches [4]–[6], [12], [22], [23]. Gao et al. [4] predict cyclists' trajectories in a car-to-cyclist scenario within a non-signalized intersection utilizing a two-step approach with a dynamic bayesian network (DBN) intention prediction and a LSTM encoder-decoder setup for position predictions based on camera data. However, they have not looked at the generalizability of their solution, and made very strong assumptions about the independence of variables. Xin et al. [5] developed an intention-aware trajectory prediction for cars in the context of autonomous driving. The authors follow a similar two-step approach, each with its own LSTM network, first forecasting the intention and using the result for the trajectory prediction. Park et al. [6] use an LSTM solution coupled with an encoder-decoder design to perform a multimodal trajectory prediction in highway scenarios. Saleh et al. [12] implement a bidirectional recurrent neural network consisting of LSTMs trained with the cyclist's position as an input. Dai et al. [22] model spatial interactions and shortcuts in their LSTM network to include car interactions. Phillips et al. [24] include interactions when predicting the behavior of human drivers at intersections. They achieve promising results by training a LSTM network to forecast whether a driver is turning left, right, or going straight ahead.

Most of these works do not focus on predicting turning behavior, or do not investigate bicycles. Furthermore, the presented solutions often base their results on very precise input data, usually collected by additionally installed sensors, utilizing camera data for accurate position data, or generating simulation data for training. Furthermore, they typically limit their analysis to one single scenario also used for training, and do not sufficiently look into the explainability of their models.

III. Bicyclist Intention Prediction at Intersections

A. Intention Prediction Model

To predict a cyclist's intention at intersections, we realize a two-layer, stacked, bidirectional LSTM network, similar to the work by Saleh et al. [12]. As shown in Figure 1, we adapt their solution to produce an intention prediction with probabilities for turning left, right, or going straight.

As an input, the sensor data at the current time t plus the last $n-1$ measurements are used (cf. Figure 1). Thereby, n defines the size of a sliding input window of measurement

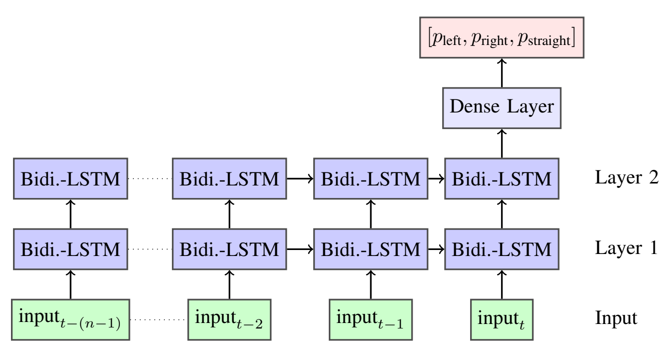


Figure 1. Bidirectional, stacked LSTM network for intention prediction with input of window size n.

Table I
INPUT FEATURES FOR INTENTION PREDICTION

Feature	Type	Description
GPS position	derived	relative GPS coordinates to intersection center
GPS accuracy	raw	estimated GPS accuracy in radius of 68 % confidence
speed	derived	based on timestamps and distances, low-pass filter
lin. acceleration	raw	indicate, if missing
rotation vector	raw	indicate, if missing

data. For each step in time, the window is moved one step. Using such a sliding input window enables our network that expects a static input size to handle traces of varying lengths. The time of crossing an intersection, and with that the number of measurement points in a trace, vary with the cyclist's speed. Hence, most traces will vary in length. For our model, we evaluate the potential sliding window sizes 1, 3, 5 and 8 with a step size of 1. As potential input features, we chose GPS position, estimated GPS accuracy, speed, linear acceleration, and rotation vector. They are summarized in Table I.

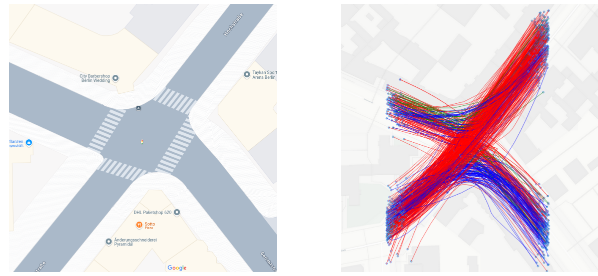
Additionally, as the measurements can include missing values, we add a masking layer to indicate missing sensor data. To prevent overfitting, dropouts are included after each bidirectional LSTM layer. In layer 1, the LSTM has 256 hidden cells. For layer 2, this value is reduced to 128 hidden cells. As our output is a multi-class classification with different probabilistic outputs that should sum to one, categorical cross-entropy loss $L(y, \hat{y})$ is applied:

$$L(y, \hat{y}) = -\sum_{i=1}^{C=3} y_i \times \log(\hat{y_i}) \tag{1}$$

with y_i being the true label, $\hat{y}_i$ being the predicted probability for class i, and C being the number of classes. We select an Adam optimizer with a default learning rate *lr* of 0.001, a momentum decay rate β_1 of 0.9, and a squared gradient decay rate β_2 of 0.9999 for its efficiency on large datasets.

B. Data Preprocessing

Our approach is trained and tested with real-world, smartphone cycling data from the SimRa dataset[1] [15]. It is part of an ongoing project allowing cyclists to record their regular rides via a smartphone app. Hence, our intention prediction requires no additional equipment, or infrastructure setup. Firstly, we read all traces collected in Berlin from June 2019 until February 2024 into a PostgresSQL database. Similar to [25], we filter out rides with multiple unrealistic position jumps ("teleportation") of more than 100 m/s, short rides of only a few seconds, and parts of rides where users stopped recording. Furthermore, we apply a Gaussian kernel filter to the GPS data to even out inaccuracies and noise. The velocities are computed based on distances between locations and time stamps. Next, a low-pass filter is applied to the velocity values smoothing the data and further reducing noise. Furthermore, as GPS positions are not as frequently recorded as IMU data, we apply linear interpolation between consecutive GPS recordings.



(a) Google maps view (b) Visualized traces

Figure 2. Selected intersection for intention prediction.

For this paper, we focus on four-street cross-shaped intersections. To achieve good training results, the dataset is combined with OpenStreetMap[3] data of Berlin, and filtered for four-street cross-intersections with more than 600 available traces. Furthermore, an intersection with a variety of different crossing behaviors is required. After analyzing multiple options, we decided for an intersection at 52.543 079 5° latitude and 13.376 406 3° longitude, as shown in Figure 2a. In total, we have 719 traces available for the chosen intersection. For further processing, traces with an estimated GPS accuracy of over 100 m were filtered out, resulting in 680 traces as illustrated in Figure 2b. In average, the GPS accuracy of all traces was 7.15 m.

C. Data Labeling and Training

Before feeding the data into the network, we labeled the traces. We choose a three-vector labeling in the format $[p_{\text{left}}, p_{\text{right}}, p_{\text{straight}}]$ where each field represents the probability p of the turning direction. The correct direction is represented with 1 while the others are set to 0. In total, 58 left turnings, 165 right turnings, and 457 straight traces are used for training, and testing. To improve generalizability of the model to other intersections, GPS positions are recomputed in relation to the intersection center, rather than using the absolute values for training. All input features are further min-max scaled to fit the $[-1, 1]$ scale of the default hyperbolic tangent activation function of LSTMs.

As the different turning behaviors for the intersection are very unevenly distributed, we apply stratified test-split on all traces before computing input windows, with a 70/30 relation for the training and testing set. The splitting is conducted before dividing the traces into windows to avoid shuffling issues of time series data. Finally, we used 41 left, 115 right, and 320 straight traces during training, and 17 left, 50 right, and 137 straight traces for testing. To balance out the uneven distribution of classes, class weights are computed. During training, we apply a batch size of 64, and limit the epochs to 500. To further prevent overfitting, we include an early stopping mechanism [26]. It stops training if the validation loss does not change significantly for at least 5 epochs, and restores the best configuration.

[3]https://www.openstreetmap.org

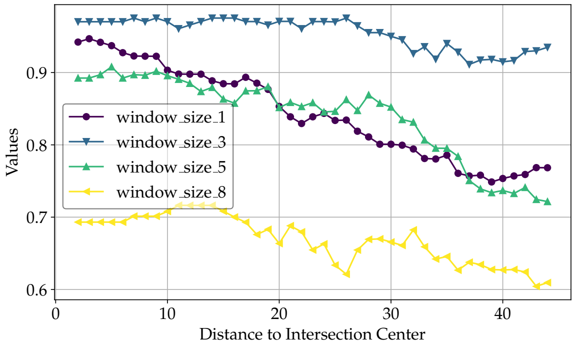


Figure 3. Weighted F1 score for different window sizes and full feature set.

IV. Optimization and Explainability Study

A. Window Optimization

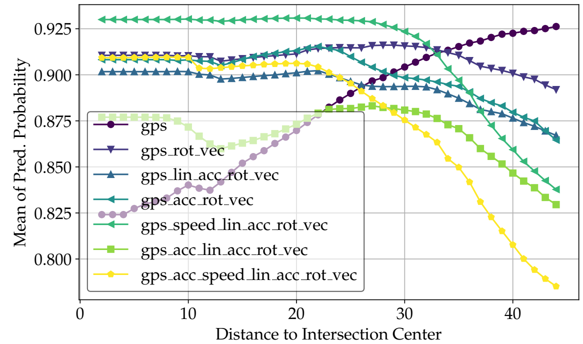
First, we train our model with the complete feature set (cf. Table I) applying four different sliding window sizes 1, 3, 5 and 8. As shown in Figure 3, the weighted F1 score gives best results for a window size of 3. Evaluating the turning intention prediction certainties, the three-window model achieves the highest probability of correctly predicting left turns, right turns, and straight-going traces. It can predict left turns with a probability of 90 % 23 m before the intersection. The intention of going straight can be forecast with a probability of 97 % 45 m before the intersection. Right turns are detected with at least 90 % 45 m before the intersection. Hence, as it has the overall best performance, we continue our explainability analysis with window size 3. Note that our results are different to findings in [12], in which a window size 5 gave the best results.

B. Feature Engineering

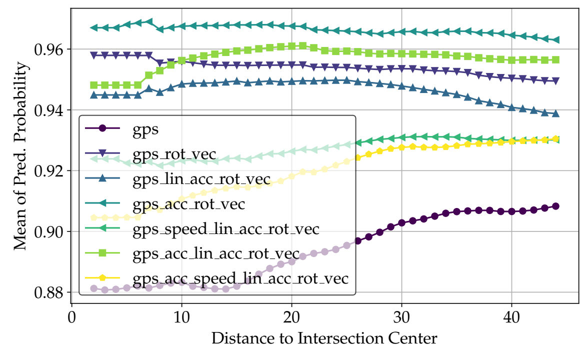
For feature engineering, we train our model with all existing subsets of features for our model. Performance results for best-performing models of each category are depicted in Figure 4.

For single-feature models, the GPS input performed best by far. Interestingly, it performs slightly better than the model with the complete input feature set for right and left turns. However, it performs significantly worse in predicting going straight. In general, all best-performing models include GPS positions. All models, except the single-feature GPS model, can predict right turns (cf. Figure 4b) and going straight (cf. Figure 4c) with a probability of at least 90 % more than 40 m before the intersection. Overall, left turns seem to be most difficult to predict. Potentially, the comparatively lower number of left turns incorporated in the training data could affect the prediction. Thus, we focus on models that perform well for left turns.

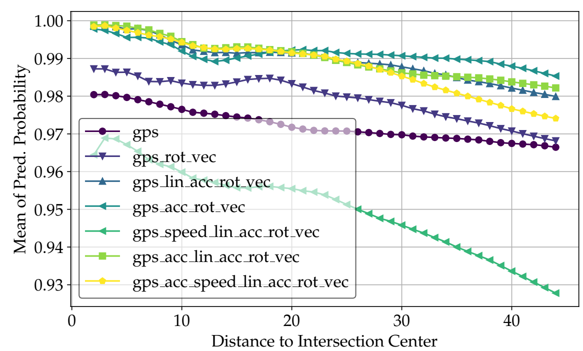
For further explainability analysis and generalizability investigations, we focus on the model trained with GPS, the estimated GPS accuracy, and the rotation vector. This model has the best weighted F1 scores and shows excellent accuracy for all intention predictions. We skip the typical step of increasing the training epochs for models with higher numbers of features as the training never continued for more than 35 epochs before



(a) Left turns



(b) Right turns



(c) Going straight

Figure 4. Results for best-performing models during feature engineering.

stopping. Furthermore, we increased the epochs for models with at least four features by doubling them. However, no significant performance improvement was achieved. Hence, we conclude that the model based on GPS, the estimated GPS accuracy, and rotation vector performs best in our evaluation. It can correctly predict left turns with at least 90 % probability already 28 m, right turns with at least 92 % probability already 45 m, and going straight with at least 97 % probability already 45 m before the intersection center.

C. SHAP Values Analysis

To evaluate each feature's contribution to predicting maneuvers, we perform a SHapley Additive exPlanations analysis. A key objective in cyclist-turn prediction is determining a cyclist's intention far enough in advance to allow other traffic participants to respond safely. We therefore focus on a distance

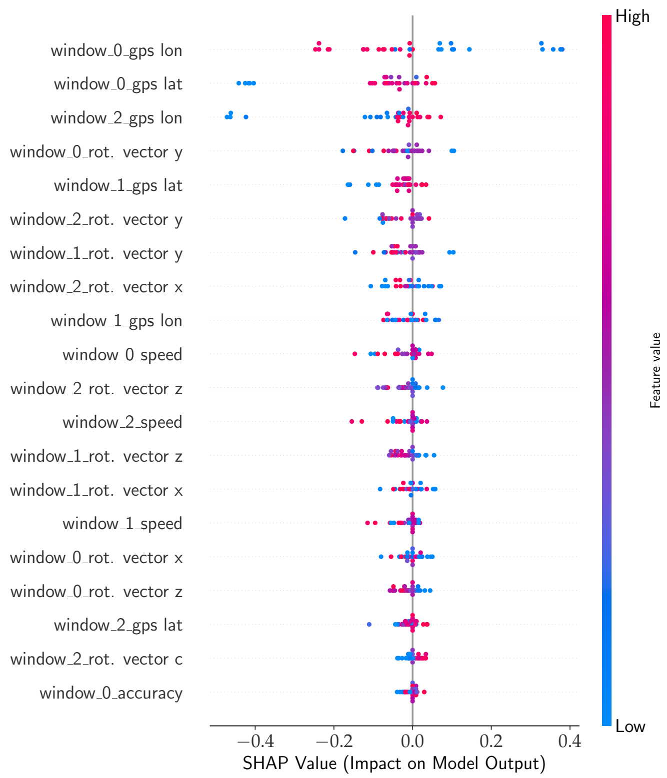


Figure 5. SHAP beeswarm plot illustrating each feature's contribution to predicting a right turn 45 m before the intersection, using the full feature set and a window size of 3.

of 45 m before the intersection. At this range, turning behaviors have often begun, yet there remains sufficient space and time for cars or nearby road users to adapt. In Figure 5, we visualize the SHAP values for correctly predicting a right turn 45 m before the intersection, given the selected input features: GPS (latitude, longitude, and estimated accuracy), speed, linear acceleration, and rotation vectors for three consecutive time windows (`window_0`, `window_1`, and `window_2`). The horizontal axis represents the SHAP value (i.e., impact on the model's output), and each dot corresponds to one sample from the test set; red dots indicate higher feature values, whereas blue dots indicate lower feature values. Features at the top of the beeswarm plot exert the greatest influence on the model's prediction (based on their mean absolute SHAP value), while those at the bottom play comparatively smaller roles.

Several informative observations can be discussed. First, GPS-related features (particularly `window_0_gps lon`, `window_0_gps lat`, and `window_2_gps lon`) appear at or near the top of the plot, underscoring that the model strongly relies on positional information when determining if a cyclist will execute a right turn well in advance of an intersection. Notably, red (i.e., higher longitude or latitude values) often correlates with positive SHAP values for right-turn classification, suggesting that specific geographic regions or directional bearings consistently increase the likelihood of predicting a right turn. Meanwhile, rotation vector components across different time windows (`window_0_rot.vector y`, `window_2_rot.vector y`, `window_1_rot.vector y`, etc.) also show substantial contributions. This is consistent with the idea that subtle orientation changes (e.g., leaning, body rotation, or handlebar movement) emerge before an overt turn and thus serve as important indicators of a forthcoming right turn. Features such as speed and GPS accuracy appear lower on the list but still contribute to the model's decisions. For example, `window_0_speed` contributes moderately—indicating that although velocity does matter, it may be secondary compared to positional shifts or orientation cues for predicting a future right turn. The estimated GPS accuracy (`window_0_accuracy`) similarly has non-negligible effects, but its impact is clearly smaller than the immediate positional and rotational information. This modest impact of GPS accuracy suggests that, for right-turn prediction at 45 m, coarse positioning data (even with varying accuracy) may be sufficient when combined with other more discriminative signals such as rotation vector changes.

Overall, the SHAP values results confirm that positional cues (GPS) and orientation dynamics (rotation vector features) are primary drivers in forecasting a cyclist's right turn. The results align with earlier observations that GPS-based features alone can provide robust hints for turn prediction, and that including orientation data further boosts the model's ability to discern turning maneuvers well in advance.

V. Evaluation

A. Performance Comparison

To evaluate the effect of memory, we compare the performance of our intention prediction to a setup where the LSTM cells are replaced by RNNs (cf. Figure 6). An equivalent number of 256 hidden cells for the first, and 128 for the second layer were used, followed by a dense layer. For a fair comparison, an equivalent sliding window approach is applied. As input, the full features, as well as the best feature combination (GPS, estimated GPS accuracy, and rotation vector) are fed into the network (cf. Table I). Accordingly, a categorical cross-entropy loss and an Adam optimizer with the same parameters are chosen. The comparison model is trained with the equivalent training-test split of 70/30 with a batch size of 64 and 500 epochs. Again, we include an early stopping mechanism to avoid overfitting. We compare this new

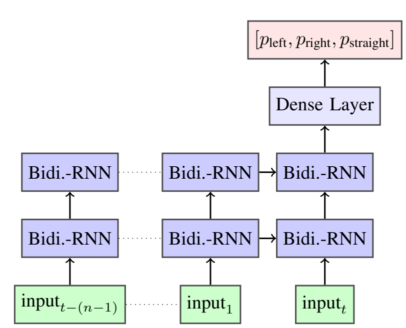
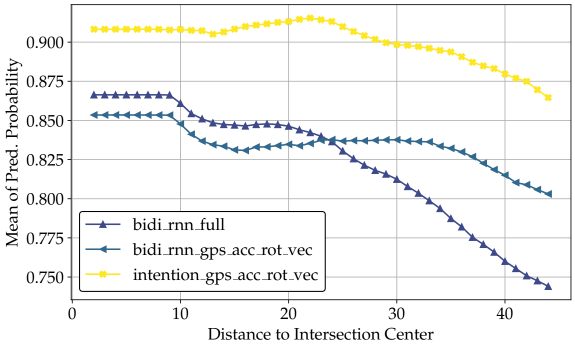
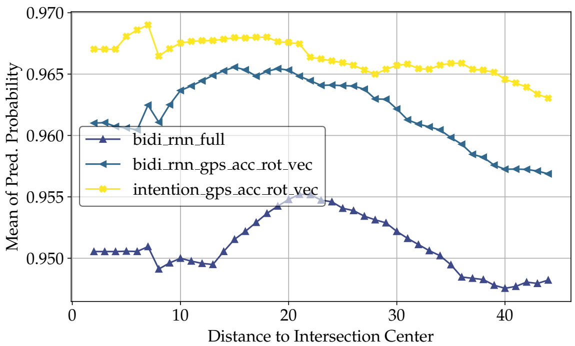


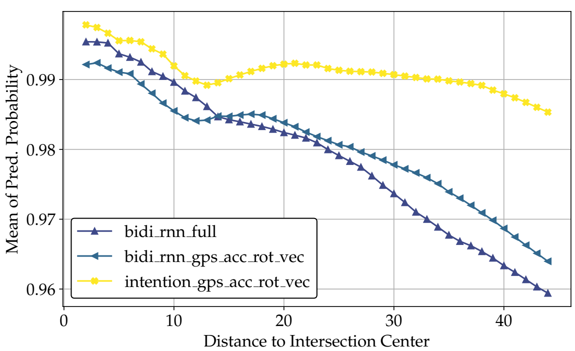
Figure 6. Comparison model based on simple RNNs.



(a) Left turns

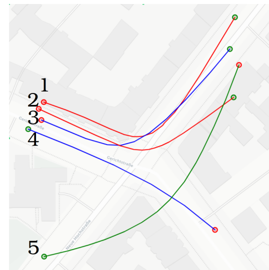


(b) Right turns

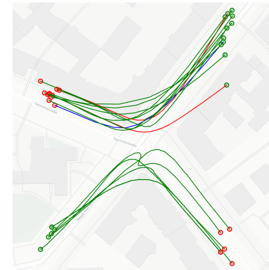


(c) Going straight

Figure 7. Performance results of RNN model and best intention prediction.



(a) Visualization of all wrongly predicted traces(red as straight, green as left, blue as right)



(b) Wrongly predicted left-turning traces(red, blue) with correctly predicted ones(green)

Figure 8. Visualization of wrongly predicted traces for the 3-feature model with GPS, GPS accuracy, rotation vector 28 m before intersection center.

setup to our best-performing three-window model with GPS data, their estimated accuracy, and the rotation vector as input.

As visualized in Figure 7, our LSTM-based intention prediction outperforms both comparison models in predicting the correct turning directions, highlighting the positive effect of memory on intention prediction. While the RNN models can still predict right turns, and going straight events with a high probability of at least 94 % 45 m before the intersection, they only achieve a minimum 85 % probability of recognizing left turns about 10–12 m before. This is at least 5 % less, and 18–20 m later compared to our approach.

B. Outlier Analysis

In addition, we manually analyzed wrongly predicted traces. As shown in Figure 8, 5 of the 204 testing traces were incorrectly predicted. The red dot marks the start of a trace, the green dot marks the end of a trace. The green trace, marked with a 5 in Figure 8a, was incorrectly identified as going left, although it goes straight through the intersection. The trace shows typical right-turning behavior in the beginning. It drifts slightly to the left, a natural movement when turning right on the bicycle, which might lead to the wrong prediction. Here, however, the trace starts turning left in the beginning which probably leads to the wrong prediction.

The two red traces, marked 1 and 2, were wrongly identified as going straight, even though they turned left. However, when visualized with other left-turning traces in Figure 8b, they continue going straight without any turning indication until reaching the intersection center.

The two blue traces, marked as 3 and 4, were classified as turning right, however, they are turning left and going straight. The first one was correctly identified as a turning process, however, for the wrong direction. When looking at Figure 8b, the trace is the most right on the street. This could lead to misclassification. Adding more left traces to the training might further improve the model.

C. Communication Requirements

Our best-performing intention prediction model with a window size of three, and GPS data, estimated GPS accuracy, and rotation vector as an input predicts all turning probabilities for up to 28 m before the intersection with a probability of at least 90 %. Hence, we can make first assumptions about how much time is left until the cyclist reaches the intersection. For each trace, we computed the average speed crossing the intersection. The traces are then split into different intention categories. We calculated the maximum average speed of all traces to identify the fastest cyclist for each maneuver. Together with the distance at which each intention can be predicted with a 90 % certainty, the time to reach the intersection center is calculated for the average, and the maximum average speed.

The results are visualized in Table II. For the fastest left-turning trace, a prediction with 90 % certainty can be made 4.98 s before the cyclist reaches the intersection center. For the fastest right-turning trace, a prediction with 90 % certainty can be made 6.02 s, and for the fastest straight-going trace, a prediction with 90 % certainty can be made 4.19 s before the cyclist reaches the center. Hence, there will be a time window

Table II
TIME ESTIMATES FOR THE TIME FROM THE MOMENT A MINIMUM OF 90 % CERTAINTY FOR CORRECT INTENTION PREDICTION TO REACHING THE INTERSECTION CENTER.

Intention	Distance	Average Speed	Max. Average Speed	Time Average Speed	Time Max. Average Speed
Left	28 m	2.54 m/s	5.62 m/s	11.02 s	4.98 s
Right	45 m	3.94 m/s	7.48 m/s	11.42 s	6.02 s
Straight	45 m	4.19 m/s	10.08 m/s	10.73 s	4.19 s

of at least 4.19 s available for sending warning messages, processing them, and reacting accordingly.

This time window is obviously more relaxed than previously anticipated warning message delays. According to the Euro NCAP recommendations [27], a forward collision warning for longitudinal scenarios should issued at least 1.7 s before a potential crash. This would result in a minimum message generating, sending, and processing time of 2.49 s for the fastest cyclist (cf. Table II). Schories et al. [28] conducted real-world experiments in a secluded environment for cyclist-car scenarios. They found that warnings as early as 4 s together with an urgent warning at 2 s could prevent 98 % potential crashes in their scenarios. However, for urban scenarios like ours, they recommend later warnings as identifying the conflict partner is difficult otherwise. Hence, assuming we need to issue a warning 3.5 s before a potential collision at the intersection center, a minimum message generating, sending, and processing time of 0.69 s for the fastest cyclist is required (cf. Table II).

D. Generalizability

1) Applicability to New Intersections: To gain first insights into the generalizability of our model, we chose an untrained but similar four-street intersection at 52.487 446 4° latitude and 13.431 550 8° longitude. The complete trace data (25 left, 33 right, and 69 straight) were applied for testing. The resulting probabilities of correct turning predictions are summarized in Figure 9. Straight-going traces continue to be the easiest to predict. However, instead of 45 m before the intersection, only 20 m predictions can be made with a certainty of 90 %. The recognition probability of left-turning

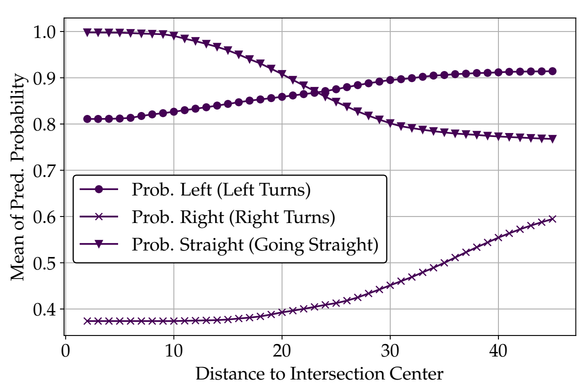


Figure 9. Prediction probabilities for intentions after converting the model to an untrained intersection.

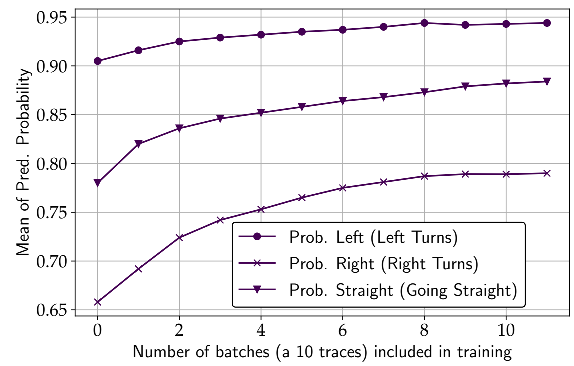


Figure 10. Prediction probabilities for intentions at 45 m while retraining the model during testing with batches of 10 traces.

traces is reduced from 90 to 80 % the closer the cyclist gets to the intersection. Multiple factors contribute to this behavior. Firstly, the intention prediction model was only trained with 58 left-turning traces. Hence, adding more traces from similar intersections to the training process could help to improve the prediction performance. Secondly, the angle between the intersecting streets of both intersections varies, resulting in a visible change in turning angle and earliness, which might affect the prediction results. Still, left turns can be predicted with a probability of at least 90 % at a distance of 45 m. For right-turning traces, the prediction probability is insufficient for application. Again, the varying angle between the intersecting streets results in a change of turning angle, and earliness. However, looking at common cyclist-car crash scenarios, right turns of bicycles are less likely to lead to accidents [29]. Overall, the model performs astonishingly well for the new intersection.

2) Online Retraining: We further evaluate the effects of online learning to see how fast our model can adjust to a new intersection. First, we test it with the full set of traces for the new intersection. Then, a batch of 10 randomly chosen traces is used for retraining the model. One batch consists of 2 left turns, 3 right turns, and 5 straight traces, representing the distribution of data available for the intersection. The resulting probabilities for the remaining test traces are then computed. This process is repeated 11 times until only 17 traces are left for evaluating the retrained model. The resulting prediction probabilities of maneuvers for a distance of 45 m are visualized in Figure 10. Even retraining with only two additional batches leads to a significant increase in correctly predicting maneuvers. Hence, either we will use continuous online retraining for improved accuracy or we will further increase the training data set with left- and right-turning traces of more intersections to generalize our trained LSTM model for other intersections.

VI. CONCLUSION

In this paper, we show the immense potential of our bidirectional, stacked LSTM intention prediction model for cyclists. After training, we demonstrate that even imprecise GPS data are sufficient to predict right turns, and straight-going traces with a certainty of 90 % 45 m, and left turns 28 m before the intersection center. The presented results show that our model can predict cyclists' intentions early using traces from noisy smartphone data. While already attaining promising first results when converted to another intersection, we aim to further enhance the performance of our model by extending the training set with additional traces from similar intersections. This requires an analysis of suitable, preferable automatic, groupings of intersections. We will extend the generalizability investigation by applying our model to multiple different groupings of intersections. Regarding intersections, we plan to explore cases with different numbers of streets, structures, and lane setups. Additionally, to increase our model's performance, we will investigate the real-time requirements of our setup, and evaluate our solution in safety scenarios sending warnings based on intention predictions. Finally, we intend to extend our model by using the intention prediction as an input for concrete bicycle trajectory forecasts.

Acknowledgments

This work has been funded by the German Federal Ministry of Education and Research (BMBF, Germany) as part of the 6G Platform under Grant 16KISK050.

References

[1] I. I. Hellman and M. Lindman, "Estimating the crash reducing effect of Advanced Driver Assistance Systems (ADAS) for vulnerable road users," *Traffic Safety Research*, vol. 4, 2023.

[2] S. Shiwakoti, S. Bikram Shahi, and P. Singh, "Deep Learning Methods for Vehicle Trajectory Prediction: A Survey," in *IoT Based Control Networks and Intelligent Systems (ICICNIS 2023)*, Singapore, Singapore: Springer, Jun. 2023, pp. 539–554.

[3] Y. Fu, C. Li, F. R. Yu, T. H. Luan, and Y. Zhang, "A Survey of Driving Safety With Sensing, Vehicular Communications, and Artificial Intelligence-Based Collision Avoidance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6142–6163, Jul. 2022.

[4] H. Gao, H. Su, Y. Cai, R. Wu, Z. Hao, Y. Xu, W. Wu, J. Wang, Z. Li, and Z. Kan, "Trajectory prediction of cyclist based on dynamic Bayesian network and long short-term memory model at unsignalized intersections," *Science China Information Sciences*, Jul. 2021.

[5] L. Xin, P. Wang, C.-Y. Chan, J. Chen, S. Eben Li, and B. Cheng, "Intention-aware Long Horizon Trajectory Prediction of Surrounding Vehicles using Dual LSTM Networks," in *21st IEEE International Conference on Intelligent Transportation Systems (ITSC 2018)*, Maui, HI, Nov. 2018, pp. 1441–1446.

[6] S. H. Park, B. D. Kim, C. M. Kang, C. C. Chung, and J. W. Choi, "Sequence-to-Sequence Prediction of Vehicle Trajectory via LSTM Encoder-Decoder Architecture," in *29th IEEE Intelligent Vehicles Symposium (IV 2018)*, Changshu City, China: IEEE, Jun. 2018, pp. 1672–1678.

[7] F. Diehl, T. Brunner, M. Le, and A. Knoll, "Graph Neural Networks for Modelling Traffic Participant Interaction," in *30th IEEE Intelligent Vehicles Symposium (IV 2019)*, Paris, France: IEEE, Jun. 2019, pp. 695–701.

[8] X. Li, X. Ying, and M. C. Chuah, "GRIP: Graph-based Interaction-aware Trajectory Prediction," in *22nd IEEE International Conference on Intelligent Transportation Systems (ITSC 2019)*, Auckland, New Zealand: IEEE, Oct. 2019, pp. 3960–3966.

[9] M. Li, T. Chen, and H. Du, "Trajectory prediction of cyclist based on spatial-temporal multi-graph network in crowded scenarios," *Electronics Letter*, pp. 97–99, Nov. 2021.

[10] D. Lee, Y. P. Kwon, S. McMains, and J. Hedrick, "Convolution neural network-based lane change intention prediction of surrounding vehicles for ACC," in *20th IEEE International Conference on Intelligent Transportation Systems (ITSC 2017)*, Yokohama, Japan: IEEE, Oct. 2017, pp. 1–6.

[11] W. Ding, J. Chen, and S. Shen, "Predicting Vehicle Behaviors Over An Extended Horizon Using Behavior Interaction Network," in *IEEE International Conference on Robotics and Automation (ICRA 2019)*, Paris, France, May 2019, pp. 8634–8640.

[12] K. Saleh, M. Hossny, and S. Nahavandi, "Cyclist Trajectory Prediction Using Bidirectional Recurrent Neural Networks," in *Advances in Artificial Intelligence (AI 2018)*, Wellington, New Zealand: Springer, Nov. 2018.

[13] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social LSTM: Human Trajectory Prediction in Crowded Spaces," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, Las Vegas, NV, Jun. 2016, pp. 961–971.

[14] K. Fragkiadaki, S. Levine, P. Felsen, and J. Malik, "Recurrent Network Models for Human Dynamics," in *IEEE International Conference on Computer Vision (ICCV 2015)*, Washington, D.C., United States: IEEE, Dec. 2015, pp. 4346–4354.

[15] A.-S. Karakaya, I.-A. Stef, K. Köhler, J. Heinovski, F. Dressler, and D. Bermbach, "Achieving Realistic Cyclist Behavior in SUMO using the SimRa Dataset," *Elsevier Computer Communications*, vol. 205, pp. 97–107, May 2023.

[16] T. Ruß and S. Naumann, "Avoiding collisions between pedestrians/cyclists and vehicles at signal controlled intersections using V2X," in *International Symposium on Future Active Safety Technology Toward zero traffic accident (FAST-zero'15)*, Gothenburg, Sweden, Sep. 2015, pp. 297–304.

[17] L. Mathuseck, J. Götz, L. Busch, and K. David, "BikeSense: Riding Behaviour Recognition Using An Instrumented Bicycle," in *22nd IEEE International Conference on Pervasive Computing and Communications (PerCom 2024)*, Biarritz, France, Mar. 2024, pp. 1–6.

[18] Z. Han, L. Xu, X. Dong, Y. Nishiyama, and K. Sezaki, "HeadMon: Head Dynamics Enabled Riding Maneuver Prediction," in *21st IEEE International Conference on Pervasive Computing and Communications (PerCom 2023)*, Atlanta, GA, Mar. 2023, pp. 22–31.

[19] S. Zernetsch, S. Kohnen, M. Goldhammer, K. Doll, and B. Sick, "Trajectory prediction of cyclists using a physical model and an artificial neural network," in *27th IEEE Intelligent Vehicles Symposium (IV 2016)*, Göteborg, Sweden, Jun. 2016, pp. 833–838.

[20] Á. Losada, A. Páez, F. Luque, L. Piovano, N. Sánchez, and M. Hidalgo, "Vehicle-to-Cyclist Collision Prediction Models by Applying Machine Learning Techniques to Virtual Reality Bicycle Simulator Data," *Applied Sciences*, vol. 14, no. 4, Feb. 2024.

[21] V. Kosaraju, A. Sadeghian, R. Martin-Martin, I. Reid, S. H. Rezatofighi, and S. Savarese, "Social-BiGAT: Multimodal Trajectory Forecasting using Bicycle-GAN and Graph Attention Networks," in *International Conference on Neural Information Processing Systems (NIPS 2019)*, Vancouver, Canada, Dec. 2019, pp. 137–146.

[22] S. Dai, L. Li, and Z. Li, "Modeling Vehicle Interactions via Modified LSTM Models for Trajectory Prediction," *IEEE Access*, vol. 7, pp. 38 287–38 296, Jan. 2019.

[23] J. P. Mercat, T. Gilles, N. E. Zoghby, G. Sandou, D. Beauvoi, and G. P. Gil, "Multi-Head Attention for Multi-Modal Joint Vehicle Motion Forecasting," in *IEEE International Conference on Robotics and Automation (ICRA 2020)*, Paris, France: IEEE, Mar. 2020, pp. 9638–9644.

[24] D. Phillips, T. Wheeler, and M. Kochenderfer, "Generalizable intention prediction of human drivers at intersections," in *28th IEEE Intelligent Vehicles Symposium (IV 2017)*, Redondo Beach, CA, Jun. 2017, pp. 1665–1670.

[25] K. Köhler, "Modeling Cyclist Behavior in SUMO based on the SimRa Dataset," Master's Thesis, School of Electrical Engineering and Computer Science, Berlin, Germany, Jul. 2021.

[26] M. K. Anam, S. Defit, H. Haviluddin, L. Efrizoni, and M. Firdaus, "Early Stopping on CNN-LSTM Development to Improve Classification Performance," *Journal of Applied Data Sciences*, vol. 5, no. 3, pp. 1175–1188, 2024.

[27] European New Car Assessment Programme (Euro NCAP), *(December 2021) Assessment Protocol - Vulnerable Road User Protection (Version 10.0.5)*, [Online]. https://cdn.euroncap.com/media/67553/euro-ncap-assessment-protocol-vru-v1005.pdf (accessed on 10.2.2025), 2021.

[28] L. Schories, N. Dahringer, U. Piram, A. Raut, S. Nikolaou, I. Gragkopoulos, I. Tsetsinas, and M. Panou, "Safety Performance Assessment via Virtual Simulation of V2X Warning Triggers to Cyclists with Models Created from Real-World Testing," *Sustainability*, Jan. 2024.

[29] M. Kuehn, T. Hummel, and A. Lang, "Cyclist-Car Accidents - Their Consequences for Cyclists and Typical Accident Scenarios," in *24th International Technical Conference on the Enhanced Safety of Vehicles (ESV 2015)*, Göteborg, Sweden: TRB, Jun. 2013.