2018 21st International Conference on Intelligent Transportation Systems (ITSC)
Maui, Hawaii, USA, November 4-7, 2018

Modeling Mixed Traffic in Shared Space Using LSTM with Probability Density Mapping

Hao Cheng* and Monika Sester*

Abstract— **In shared spaces, vulnerable road users (i.e., pedestrians) are encouraged to directly interact with other road users (i.e., vehicles). They self-organize to give or take right-of-way without being regulated by explicit traffic rules. The safety based on their behavior patterns in such areas is, hence, critical and needs to be fully investigated. In this paper, we first carry out a collision probability method based on safety distance to quantitatively study behavior patterns in mixed traffic. Then we propose a Long Short-Term Memories recurrent neural networks model that takes 2D trajectory data at discrete time steps and incorporates collision probability that captures user behavior patterns as a density mapping function for trajectory prediction. The model handles collisions explicitly based on the relative positions of the neighboring users in an ego user's vicinity with considering the impact of personal space and vehicle geometry. It also provides a friend flock detection mechanism to allow close interactions between friends. After training by real–world trajectories, the model outputs comparative results to the state–of–the–art methods for predicting three-second trajectories in complicated situations in a shared space. It can be applied for intent detection and on-board alarming system for autonomous driving when interacting with multimodal road users in shared spaces.**

I. Introduction

In shared spaces, road signs, signals, and markings are removed to allow mixed traffic directly interact with each other [1]. The traffic engineer Reid defined it as a street encouraging pedestrian movement and reducing the dominance of vehicles without explicit traffic rules [2]. The lack of regulations makes interactions between multimodal road users more complex compared with conventional designs. All users have to follow informal social protocols and negotiation to use the road resources, and avoid any potential collisions. Understanding their behavior and predicting their intents after a very short observation time are crucial to traffic planning and autonomous driving in such areas.

The established models (e.g., Social Force [3]) for pedestrian locomotion may not be sufficient for modeling mixed traffic in shared spaces given various impacts. For example, heterogeneity of transport modes [4], dynamic environments [5], and demographic attributes (e.g., age, gender, time pressure [6]) of road users can lead to distinct results for trajectory choice. Hence, predicting individual behavior with the expectation of neighboring users in his or her vicinity is of great challenge. In the recent research, Deo and Trivedi proposed an extended Variational Gaussian Mixture model based on three-second observed trajectories for learning and predicting pedestrian behavior around vehicles in an uncontrolled intersection [7]. But their model does not provide a mechanism to handle direct vehicle-pedestrian interactions. There are models which deal with mixed traffic. For example, the extended Social Force model [4], [8]–[10] uses a multi-layer mechanism to manually classify interactions according to their transport mode (e.g., pedestrian vs. pedestrian or pedestrian vs. car), then handles the interactions with predefined force rules. The Game-based model proposed in [11] reduces the complexity of interactions by constraining the type of game, the number of players, and the mechanism of cooperation between players. When applying these models to real–world scenarios in shared spaces, they still suffer the lack of scalability and require manual setups and fine-tuning.

As the success of deep learning techniques and the availability of large–scale datasets, neural networks–based models, especially Long Short-Term Memories (LSTM) [12] have been applied for pedestrian behavior [13] and car-following [14] modelings. They are used to extract latent features of trajectories in forward propagation and optimized for the prediction of the next positions in time during back propagation.

However, the motion of an ego user in the near future is not merely decided by his or her past movement and the desired destination, but also the expectation of and the interaction with the neighboring users and environments. Helbing and Molnar call the impacts attractive and repulsive effects in Social Force model [3]. To extend LSTM for this problem, a pooling layer that processes a centralized bounding grid of the ego user was introduced by Alahi et al. to capture the interactions in the ego user's vicinity (Social-LSTM). It outperforms rule-based models (i.e., Social Force [15] and Iterative Gaussian Process [16]) substantially. Nevertheless, there was no explicit explanation about how road users perceive collision risks or statistical evidence showing how the size of the bounding grid should be defined. While the size of interactive zone may vary across multimodes in mixed traffic. Social-LSTM only deals with homogeneous road users, it is not sufficient to be directly applied for mixed traffic modeling. It also did not consider the personal space of pedestrians, which significantly impact their perceptions of safety distance in interactions at different positions [17]. Moreover, in general, the closer the road users to each other, the higher probability of collisions, but which may not be the case for friends (within a pedestrian group). Furthermore, given the relatively large area an individual vehicle occupies, the geometry of vehicles also need to be considered when vehicles are involved in interactions.

*Institute of Cartography and Geoinformatics, Leibniz Universität Hannover, Germany hao.cheng@ikg.uni-hannover.de, monika.sester@ikg.uni-hannover.de

978-1-7281-0323-5/18/$31.00 ©2018 IEEE

In this paper, we first statistically analyze the perceptions of collision risks based on safety distance across multimodal road users—pedestrians and vehicles—in a shared space using real–world trajectory data. This is needed to elicit the interactions among them, which considers pedestrian personal space and vehicle geometry. Meanwhile, the collision probability of the friends for an ego pedestrian is cancelled out using a flock detection function [18]. Then we propose an LSTM-PDM model which takes the 2D coordinates of the ego user and the interactive collision probabilities of the ego and the neighboring users, called probability density mapping (PDM), as a concatenated input to predict the ego user's trajectory in the next three seconds.

II. Dataset

Real–world mixed trajectories in a shared space close to a busy train station in the German city of Hamburg were collected. As denoted by Fig. 1, a 63-meter long street is shared by vehicles, cyclists, and pedestrians. A maximum speed $20\,\mathrm{km/h}$ and a priority over other types of road users are granted to vehicles. Pedestrians and cyclists can cross the street at any point from both sides. However, as showed by the recorded video, rather than strictly following the traffic regulation, vehicles, cyclists, and pedestrians negotiated the space spontaneously and often gave priority to each other. In a 30-minute video, there were 338 vehicle, 1115 pedestrian, and only 22 cyclist trajectories. To avoid biases caused by the small sample size, in this paper, we do not include cyclists for our study, but only focus on pedestrians and vehicles.

For data pre-processing, the video ($30\,\mathrm{fps}$) was firstly decomposed into single frames followed by a distortion removal. Then every 15th frame was selected for extracting the trajectories using the tool tracker[1]. Each trajectory contains the 2D coordinates measured by the location of the mass at discrete time steps with each time step lasting 0.5 second for the given user. The first ten minutes ($31\,\%$ of the dataset) are saved as a test set and the rest are used for training [10], [19].

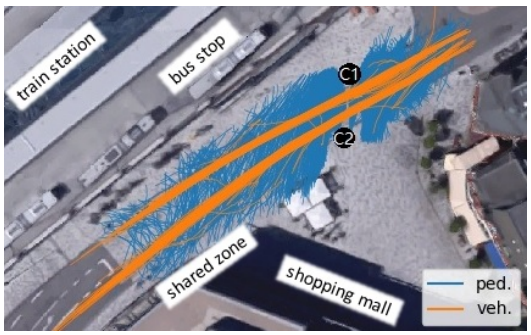


Fig. 1. Accumulated trajectories collected by two static cameras (C1 and C2) with an elevation seven meters facing two directions in a shared space. (Background image: Imagery ©2017 Google, Map data ©2017 GeoBasis–DE/BKG(©2009), Google)

[1]https://physlets.org/tracker

III. Analysis of User Behavior

A. Collision Probability

Similar to the approach Hacohen et al. used for pedestrian crossing in congested traffic [20], we map the probability of collision between two moving road users based on the minimum safety distance they maintain and the corresponding angle between the ego and the neighboring users. Instead of using a Gaussian distribution, here we follow the idea of the repulsive potentials with an exponential function in Social Force model [3]; the probability of collision is implemented as a negative exponential relation to the distance.

In real–world situations, collision probability is also impacted by personal space and vehicle geometry. An elliptic protective zone is more likely maintained by pedestrians around their body such that awareness of collisions in front is stronger than from the side [17]. The mass of a vehicle may not accurately represents its position given its size—approaching to any border of the vehicle may lead to collisions. Therefore, the collision probability increases exponentially with a moving user approaching another user's personal space or vehicle geometry until they collide into each other by touching the border.

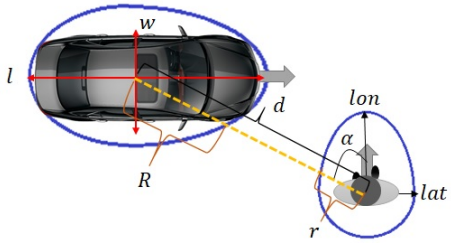


Fig. 2. Personal space and vehicle geometry. Here α is defined regarding the pedestrian as an ego user and the vehicle as a neighboring user.

Fig. 2 depicts the approximate ellipse-like shapes for the pedestrian personal space and the vehicle geometry (denoted by navy color) with respective radii r and R. α stands for the angle between the heading of the ego pedestrian and the position of the neighboring user (here the vehicle). lon and lat stands for the pedestrain's longitudinal and lateral distances for the personal space respectively. Equation (1) computes the radius of the personal space at the given α.

$$r = \begin{cases} \sqrt{\mathrm{lon}^2(\cos\alpha)^2 + \mathrm{lat}^2(\sin\alpha)^2} & \text{if } \alpha \in [-90°, 90°] \\ \mathrm{lat} & \text{else.} \end{cases} \tag{1}$$

In the shared space we analyze, personal space is squeezed by busy traffic flows. $0.63\,\mathrm{m}$ and $0.39\,\mathrm{m}$ are taken for lon and lat respectively based on the lower bound of untouched ranges to most pedestrians (over $95\,\%$), which are smaller than the values found in a laboratory [17].

Similarly, Equation (2) depicts the approximate geometry of the given vehicle. l and w stand for the approximate length and width of the ego vehicle. $3\,\mathrm{m}$ and $1.6\,\mathrm{m}$ are taken respectively. α is defined in the same way as above but now for the ego vehicle. γ is the parameter to locate the mass

along the approximate length of the vehicle, 0.6 is take in our study.

$$R = \begin{cases} \sqrt{(\gamma * l)^2(\cos\alpha)^2 + (\frac{1}{2}w)^2(\sin\alpha)^2} & \text{if } \alpha \in [-90°, 90°] \\ \sqrt{[(1-\gamma) * l]^2(\cos\alpha)^2 + (\frac{1}{2}w)^2(\sin\alpha)^2} & \text{else.} \end{cases} \tag{2}$$

Equation (3) calculates the collision probability using the probability density function with an exponential distribution for pedestrian ego users. d stands for the distance between the ego and the neighboring users (see Fig. 2). The rate parameter λ in the exponential formula is set to one. For calculating $p(d|\text{veh})$ of vehicle ego users, r is substituted by R in Equation (3). When the neighboring user's position is inside the ego user's personal space or vehicle geometry, the value is set to one. The collision probability can be treated as an indicator for the model (see Section IV-A) to avoid collision in prediction time.

$$p(d|\text{ped}) = \begin{cases} \lambda e^{-\lambda(d-r)} & \text{if } d - r > 0 \\ 1 & \text{else.} \end{cases} \tag{3}$$

B. Pedestrian Flock

To cancel out erroneous collision probability caused by friends of the ego pedestrian—small safety distance does not necessarily mean high collision probability for friends—flock detection function [18] is used to detect friends. A pedestrian flock F is define by Equation (4).

$$F \leftarrow \text{if } p(d) \geq 0.1 \text{ and } \frac{t}{\tau} \geq 75\,\%, \text{ where } t \geq 4. \tag{4}$$

t refers to the duration that a neighboring user stays close to the ego user. τ stands for the total time steps of the trajectory for the ego user. A neighboring user is considered to be a friend of the ego user if the neighboring user tolerates collision probability that is equal or higher than 0.1 and walks with the ego user during at least $75\,\%$ of the total time steps. The lower bound of t is to guarantee that at least 4 time steps need to be observed.

The pedestrian flock function is used to detect pedestrian groups among the 1115 pedestrians in our dataset. Only about $40\,\%$ pedestrians walked alone and $60\,\%$ walked in groups, which is similar to the findings by Moussaïd et al. that up to $70\,\%$ pedestrians in a commercial street are moving in groups [21]. Fig. 3 depicts the percentages of seven different sized pedestrian groups (including the ego user) with a descending trend with group size, which is in line with the results in [21].

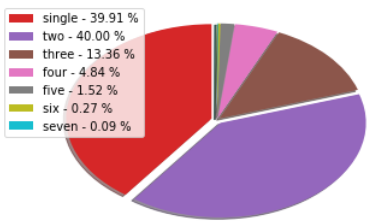


Fig. 3. The percentages of pedestrians in different sized groups

C. Behavior patterns

In order to avoid collisions in shared space, a minimum distance at travel time must be maintained as a safety distance by road users. However, the safety distance may be perceived significantly differently by multimodal road users in heterogeneous interactions they encounter. We first retrieve the minimum distances a user maintains encountering with other users in their coexistence time. Then we use the collision probability defined above with the consideration of personal space and vehicle geometry to quantitatively study the impact of the length of safety distance on interactions.

Firstly, we estimate whether the size of pedestrian groups will make a difference for safety distance a single road user maintains. Interestingly, the size does not significantly impact the safety distance an ego user (i.e. pedestrian or vehicle) may keep. Moreover, the standard deviations extend to a relatively large range, which means that the safety distance can vary largely for different ego users even interacting with pedestrian groups having a same size. However, the overall safety distances maintained by the vehicle ego users are larger than the ones maintained by pedestrian ego users across all different sized groups (see Fig. 4).

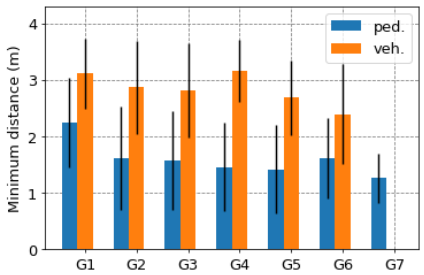


Fig. 4. The mean minimum distance from pedestrian/vehicle ego users to pedestrian groups with different sizes

This above pattern indicates that vehicle ego users have a strong perception of collision risks over others and are prudent when interacting with vulnerable road users, which is in accordance with the findings by Rinke et al. [4]. This brings us to further analyze how different types of ego users tolerate collision risks when interacting with others in the shared space. It is worth mentioning that in the following analysis, the impact of friends on pedestrians are excluded. Table I denotes the mean and median minimum distances in the interactions between different types of road users. When the ego user type is pedestrian, the ego users maintain a significantly smaller safety distance interacting with pedestrian neighboring users than vehicle neighboring users. Similarly, when the ego user type is vehicle, the ego users maintain a significantly smaller safety distance when interacting with pedestrian neighboring users than vehicle neighboring users. Table II lists the corresponding statistics for MannWhitney U test. Furthermore, the mean and median minimum distances for vehicle ego users are larger than pedestrian ego users when interacting with pedestrian and vehicle neighboring users respectively.

TABLE I. Mean and median minimum distances in the interactions between different types of road users. The measurement unit is meter

Ego	Neighbor	Mean distance	Median distance	Std.
ped.	ped.	1.98	1.99	0.80
	veh.	2.57	2.64	0.61
veh.	ped.	2.86	2.89	0.76
	veh.	3.28	3.28	0.41

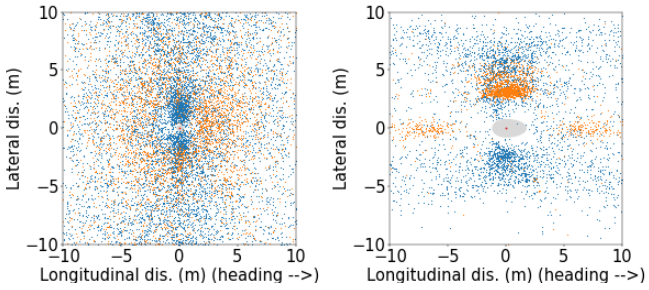
TABLE II. Comparison of minimum distances in the interactions between different types of road users using MannWhitney U test (alternative = less)

Ego	Neighbors	U-statistic	Rank-biserial r-value	p-value
ped.	ped. / veh.	517335.5	0.43	0.01***
veh.	ped. / veh.	330826.0	0.35	0.01***

However, with large standard deviations, relative angle α has less impact on road user behavior than safety distance d (see Fig. 2). The interactions at minimum distances scatter around the ego users in pedestrian vs. pedestrian[2], pedestrian vs. vehicle, and vehicle vs. pedestrian. Nevertheless, most of the interactions at minimum distances concentrate on the left side of the vehicle ego users in vehicle vs. vehicle, which means that vehicles tend to keep a larger longitudinal distance than lateral distance when the lanes in different directions are very close. Fig. 5 visualize the distributions of interactions at minimum distances. It intuitively shows that the personal space and the vehicle geometry, and also the distributions of interactions at minimum distances are distinctive between pedestrian and vehicle ego users.

TABLE III. Mean and median values of α at the corresponding minimum distances in the interactions between different types of road users. The measurement unit is degree (-180°, 180°), and the right side is negative and the left side positive

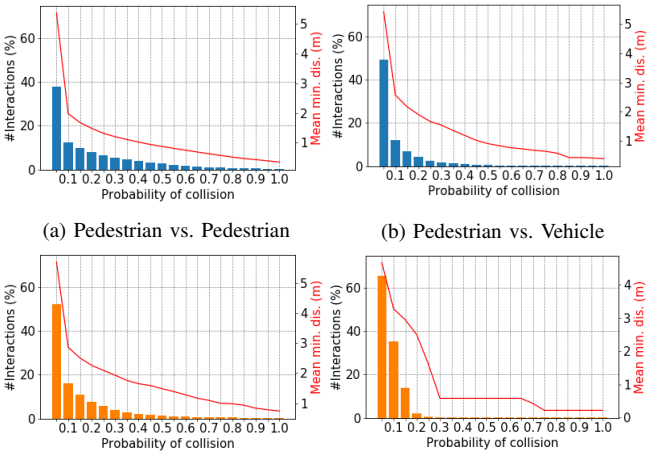
Ego	Neighbor	Mean angle	Median angle	Std.
ped.	ped.	8.09	41.68	97.78
	veh.	-6.02	-3.12	83.07
veh.	ped.	-22.36	-64.92	92.41
	veh.	82.80	87.73	37.98



(a) Pedestrian (b) Vehicle

Fig. 5. Accumulated distribution of neighbors in the vicinity of ego users. (a) the personal space and (b) the vehicle geometry are denoted by an approximate grey shadow in the middle respectively. The mass of ego users are denoted by a red dot in the middle. Due to errors in data acquisition and small-size vehicles, few neighbouring users interact with ego users in the personal space or the vehicle geometry shadow

Instead of using the length of safety distance to analyze the level of risks pedestrians and vehicles could bear, we convert it to risk probability using Equation (3) for more intuitive understanding. we segment the collision probability [0, 1] to 20 equal bins representing different risk levels, and we compare the numbers of interactions relative to the collision probability levels. Fig. 6 posts the percentages of interactions at discrete collision probability levels and the corresponding safety distances. In Fig. 6(a) and (b), as the collision probability level increases, the number of interactions in pedestrian vs. vehicle decreases faster than pedestrian vs. pedestrian. In Fig. 6(c) and (d), the descending trend is more profound for vehicle vs. vehicle than vehicle vs. pedestrian. Over $50\,\%$ of the interactions in vehicle vs. pedestrian happen at a collision probability of 0.1 with a safety distance about $5.8\,\mathrm{m}$. But most interactions in vehicle vs. vehicle can only tolerate a collision probability up to 0.2 with a safety distance at least $2.5\,\mathrm{m}$. These findings demonstrate that vehicle drivers are more cautious in general than pedestrians, especially when they encounter with other vehicles.



(a) Pedestrian vs. Pedestrian (b) Pedestrian vs. Vehicle

(c) Vehicle vs. Pedestrian (d) Vehicle vs. Vehicle

Fig. 6. Number of interactions at mean minimum distance relative to the collision probability for the interactions between different types of road users. (a) - (b) pedestrian vs. pedestrian and vehicle, respectively; (c) - (d) vehicle vs. pedestrian and vehicle, respectively

IV. Trajectory Prediction

A. LSTM-PDM Model

In this section, we propose an LSTM-PDM model to predict trajectories of mixed traffic in a shared space, which takes the initial steps of the ego user and the interactive collision probabilities of the ego and neighboring users as a concatenated input to predict the ego user's trajectory in the near future. It is motivated by Social-LSTM [13]. However, instead of using a centralized bounding grid with a fixed size to pool the existence of neighboring users in the ego user's vicinity, we use collision probability to process interactions. As showed in Section III, it captures behavior patterns of multimodal road users, including the information of the heading of the ego user and the relative position of the neighboring user, as well as the impact of personal space and vehicle geometry. This also provides the

[2]the left of *vs.* is ego user and the right is neighbouring user

model an indicator with a probability to explicitly handle collision risks in prediction time. Meanwhile, the collision of friends, which is very unlikely in reality, is set to zero by the flock function. We assume that the perceptions of risks for the ego and the neighboring users are independent since no explicit communication between them are established. Therefore, we define PDM (probability density mapping) as convolutional collision probabilities perceived by the ego and the neighboring users, denoted by Equation (5).

$$\text{PDM} = \begin{cases} P(e) * P(n) & \text{if } (e,n) \notin F \\ 0 & \text{else,} \end{cases} \tag{5}$$

where $P(e)$ and $P(n)$ stands for the respective collision probability for the ego and neighboring users, and calculated by Equation (3) with the information of the transport mode. If the ego and the neighboring users belong to the same flock F, then PDM value is set to zero. It means that the collision probability between friends is zero.

Every user is assigned to a single LSTM, which is fed with a spatial part to store the 2D coordinates of the ego user and a tensor part to process the convolutional probability PDM. Those two parts are embedded separately with Rectified Linear Unit (ReLU) as depicted by Equation (6) before concatenation. W_S and W_T stand for the embedding weights for the spatial input and the tensor input respectively. t refers to the time step and i is the index for the given ego user.

$$S_t^i = \text{ReLU}(W_S \cdot (x_t^i, y_t^i)); \quad T_t^i = \text{ReLU}(W_T \cdot PDM_t^i). \tag{6}$$

The embedded spatial input S_t^i and the embedded tensor input T_t^i are concatenated to form a complete input for the LSTM cell. Equation (7) denotes the forward propagation. h_{t-1}^i is the hidden state at time $t-1$ and W stands for the corresponding weights for the LSTM.

$$h_t^i = \text{LSTM}[h_{t-1}^i, (S_t^i + T_t^i), W]. \tag{7}$$

The model is trained using a similar method as [13] and [19], which was initially proposed by [12]. The output of the neural networks is a 5D vector to parameterize a bivariate Gaussian distribution (see Equation (8)) at time t, including a 2D vector μ^i for the arithmetic means of the respective distributions in x and y coordinates. a 2D vector σ^i for the corresponding standard deviations, and ρ^i the correlation. In turn, the Gaussian distribution is used to predict the position of user i at the next time step $t+1$.

$$(\hat{x}^i, \hat{y}^i)_{t+1} \sim \mathcal{N}(\mu^i, \sigma^i, \rho^i)_t. \tag{8}$$

A negative log–likelihood loss function is used to calculate the cost of the predicted trajectories against the true trajectories using Equation (8) and Equation (9), It sums up the complete loss for the given user at all predicted time steps.

$$\text{Loss} = -\sum \log \Pr(x_{t+1}^i, y_{t+1}^i | \mu_t^i, \sigma_t^i, \rho_t^i). \tag{9}$$

Furthermore, least square errors (L2) are applied as a regularization to penalize all the weights (i.e., W_S, W_T, W) in the model to reduce overfitting. The overall error loss, thereby, is the sum of the log–likelihood loss and L2, and is optimized employing Stochastic Gradient Decedent.

B. Evaluation Metrics and Baselines

We use the same evaluation metrics as the authors did in [19]—average Euclidean and Hausdorff [22] distances, average speed deviation, and average heading error—to measure the errors between the predicted and true trajectories. Average Euclidean distance calculates the pointwise displacement error from a predicted position to the true position aligned in time horizon. However, Hausdorff distance computes the largest distance from a predicted trajectory to the true trajectory without time step alignment. This measurement relaxes on errors caused by speed difference but emphasizes on overall displacement. Akin to Euclidean distance, both speed deviation and heading error are computed at each time step. More details about evaluation metrics can be found in [19].

In order to estimate the performance for LSTM-PDM in comparison with other models, a linear model (simple Kalman Filter) and Social-LSTM [13] are treated as baselines. The simple Kalman Filter linear regression model does not consider any interaction or transport mode for all types of road users in the shared space. As mentioned before, Social-LSTM model considers interactions between road users, but it treats all types of road users the same. We also compare LSTM-PDM with the recently extended Social-LSTM model named ViewType-LSTM [19], which incorporates sight of view of ego users and pairwise–aware interactions to differentiate the involved user types.

C. Results

The same as Social-LSTM and ViewType-LSTM, LSTM-PDM observes six time steps (three-second trajectories) and predicts the trajectories in the next six time steps. Since the estimated state is updated by the observation once available in the Kalman Filter model [23], it does not need to observe six time steps. Hence, the performance of the Kalman Filter model is measured directly in prediction time. Table IV lists the results.

TABLE IV. Prediction errors for all the models. Euclidean and Hausdorff distances are measured by meter, speed deviation is measured by meter per second, and heading error is measured by degree. The best values are highlighted in boldface

Meas. metrics	User type	Kalman Filter	Social LSTM	ViewType LSTM	LSTM PDM
Avg. Euclidean distance (m)	ped.	2.52	0.77	0.71	**0.68**
	veh.	2.51	1.15	**1.02**	1.03
Avg. Hausdorff distance (m)	ped.	7.94	1.24	**1.08**	1.10
	veh.	6.87	1.48	**1.26**	1.47
Avg. speed deviation (m/s)	ped.	4.14	0.17	0.17	**0.16**
	veh.	3.95	0.41	**0.37**	0.39
Avg. heading error (°)	ped.	**13.84**	36.79	31.79	35.17
	veh.	**12.74**	26.39	22.28	19.62

D. Discussion

The Kalman Filter model fits a line that tries to capture the linear momentum of the trajectories, particularly with the smallest heading errors for prediction compared with the other models. This is because most of the trajectories, mainly vehicle trajectories, in such a shared space tend

to be linear as guided by the space layout (see Fig. 1). However, its drawback is largely showed by the failure to simulate the behavior of interactions for road users, which are denoted by the large errors measured by average Euclidean and especially Hausdorff distances, and average speed deviation. The momentum of road users are often interrupted by interactions; they adjust their locomotion and speed (acceleration and deceleration) to avoid collisions in the interactions with other users.

In comparison with Kalman Filter, Social-LSTM adds a pooling layer to signalize the existence of the neighboring users in the given ego user's vicinity. Thus, interactions between them can be captured by this pooling layer [13]. This has been proven by the largely reduced errors measured by average Euclidean and Hausdorff distances, and average speed deviation (see the third and forth columns in Table IV).

The extended Social-LSTM model ViewType-LSTM incorporates user type by feeding the model with user's sight of view and pairwise–aware interactions [19]. For example, pedestrian has a truncated interactive grid according to his or her sight of view, and pedestrian vs. vehicle is treated differently than vehicle vs. pedestrian. The prediction errors have been further reduced compared with Social-LSTM (see the forth and fifth columns in Table IV).

Different from ViewType-LSTM that uses a heuristic mechanism to mimic how road users see and react to others, LSTM-PDM has an explicit mechanism—probability density mapping—to quantify the level of risks of collisions, as well as a flock function to relax on close interactions between friends. It also considers the impact of personal space and vehicle geometry. The performance for LSTM-PDM is remarkably better than the two baselines and very close to the performance for ViewType-LSTM (see the fifth and sixth columns in table IV). In some regards, LSTM-PDM even slightly outperforms ViewType-LSTM. For example, the errors measured by average Euclidean distance and average speed deviation are smaller for LSTM-PDM in comparison with ViewType-LSTM for pedestrians. This might be explained by the contribution of the flock function that treats pedestrian friends as a complete flock with zero collision risk. In addition, the average heading error for vehicles decreased by $12\,\%$.

E. Qualitative Study

We carry out a qualitative study to demonstrate the potential of LSTM-PDM for intent detection and on-board alarming system for autonomous driving when interacting with multimodal road users in shared spaces. LSTM-PDM is tested on real–world mixed traffic trajectory prediction. In this context, Fig. 7 showcases three complicated situations in interactions between multiple road users.True trajectories are denoted by dot lines with black color and predicted ones are denoted by color coded dot lines. Blue color represents pedestrians and orange vehicles. A dot with a larger size represents a latter time step.

In Fig. 7(a), after observing six time steps, LSTM-PDM is capable to predict the direction and speed for vehicles with several pedestrians crossing the street—three predicted vehicle trajectories overlay their corresponding true trajectories in both directions. On the lower left corner, a pedestrian lingering and waiting for the oncoming vehicle from right to left is correctly predicted. LSTM-PDM also recommends reasonable trajectories for the other pedestrians.

In the middle of Fig. 7(b), a pedestrian starts to cross the street from down to up when the safety distances between two leaving vehicles are large enough. However, the predicted trajectory has a feasibly smaller turning angle than the true trajectory for the pedestrian. In this case, the pedestrian does not necessarily need to adjust the orientation as large as he or she did in reality since the front space is free for him or her.

In Fig. 7(c), a standby for a pedestrian waiting to cross the street is predicted by LSTM-PDM, and a deceleration of the oncoming vehicle from the left to the right is also recommended by the model to avoid collision with the pedestrian. For the other three vehicles, interestingly, LSTM-PDM suggests faster leaving speed for them as the safety distances are large—over 15 meters. They are barely interact with each other.

Whereas, in this empirical study, LSTM-PDM has only been tested in an area with a shared street. Most of the time, vehicles have to stay on the clearly defined street, which can largely limit their movement. Besides, the number of observed cyclists are not enough to be included for the study. These can be circumvented by training the model with improved datasets that have divergent space layouts and more balanced multimodal road users in future work.

V. Conclusion and Future Work

In this paper, we first statistically analyze how multimodal road users perceive collision risks based on safety distance, and then we incorporate PDM derived from the statistical analyses into LSTM for mixed traffic trajectory prediction in a shared space. The proposed model explicitly handles collisions by a probability that considers the impact of personal space and vehicle geometry, and it also relaxes on the close interactions between friends. Our model predicts rational and feasible trajectories based on a short-time observation with collision avoidance. Instead of manual setups and fine tuning, it can be trained by real–world trajectory data. The model has been demonstrated on real–world scenarios to showcase the potential applicability, e.g., intent detection and on board alarming system for autonomous driving in shared spaces.

In future work, shared spaces with divergent layouts (e.g., roundabout, intersection) and more transport modes will be studied. Moreover, in this study we do not consider the context– and design–specific factors [6] (e.g., curbside, nearby facilities) which may also influence road user behavior in shared spaces. Thus, these types of contexts will be leveraged in our future work to further boost the performance for mixed traffic trajectory prediction.

Acknowledgments

The authors cordially thank the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foun-

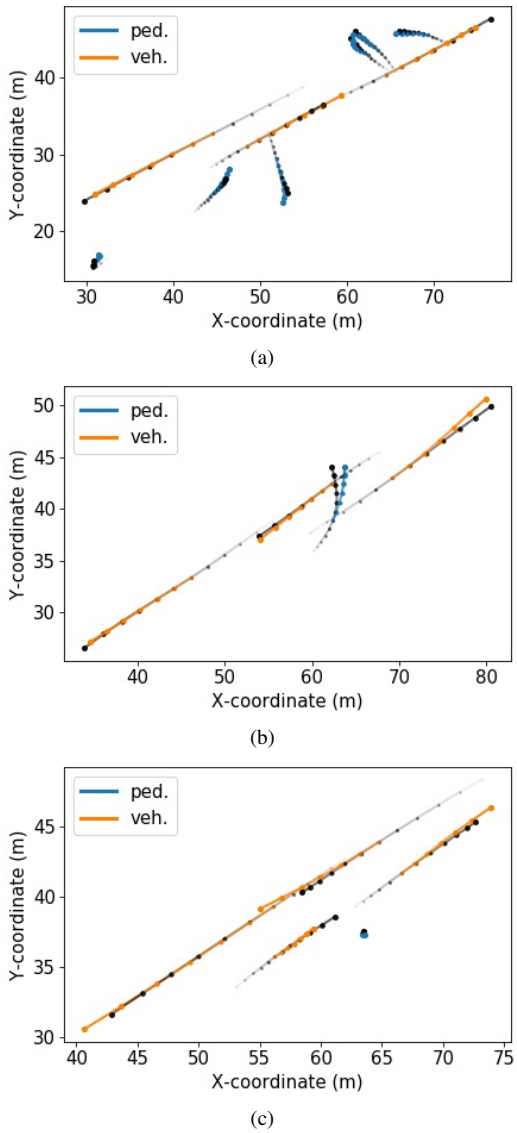


Fig. 7. Predicting trajectories for the next six time steps after observing six time steps. The true trajectories are in black color and the predicted ones are color coded—pedestrians are in blue and vehicles in orange. A marker with larger size and opacity denotes a later time point

dation) - 227198829/GRK1931 and the research project MODIS (Multi mODal Intersection Simulation) for providing the dataset of road user trajectories used in this work.

References

[1] B. Hamilton-Baillie, "Shared space: Reconciling people, places and traffic," *Built environment*, vol. 34, no. 2, pp. 161–181, 2008.

[2] S. Reid, *DfT Shared Space Project Stage 1: Appraisal of Shared Space*. MVA Consultancy, 2009.

[3] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Physical review E*, vol. 51, no. 5, p. 4282, 1995.

[4] N. Rinke, C. Schiermeyer, F. Pascucci, V. Berkhahn, and B. Friedrich, "A multi-layer social force approach to model interactions in shared spaces using collision prediction," *Transportation Research Procedia*, vol. 25, pp. 1249–1267, 2017.

[5] G. Andrienko, N. Andrienko, P. Bak, D. Keim, S. Kisilevich, and S. Wrobel, "A conceptual framework and taxonomy of techniques for analyzing movement," *Journal of Visual Languages & Computing*, vol. 22, no. 3, pp. 213–232, 2011.

[6] I. Kaparias, M. G. Bell, A. Miri, C. Chan, and B. Mount, "Analysing the perceptions of pedestrians and drivers to shared space," *Transportation research part F: traffic psychology and behaviour*, vol. 15, no. 3, pp. 297–310, 2012.

[7] N. Deo and M. M. Trivedi, "Learning and predicting on-road pedestrian behavior around vehicles," in *Intelligent Transportation Systems (ITSC), 2017 IEEE 20th International Conference on*. IEEE, 2017, pp. 1–6.

[8] F. Pascucci, N. Rinke, C. Schiermeyer, B. Friedrich, and V. Berkhahn, "Modeling of shared space with multi-modal traffic using a multi-layer social force approach," *Transportation Research Procedia*, vol. 10, pp. 316–326, 2015.

[9] C. Schiermeyer, F. Pascucci, N. Rinke, V. Berkhahn, and B. Friedrich, "A genetic algorithm approach for the calibration of a social force based model for shared spaces," *Proceedings of the 8th International Conference on Pedestrian and Evacuation Dynamics (PED)*, 2016.

[10] F. Pascucci, N. Rinke, C. Schiermeyer, V. Berkhahn, and B. Friedrich, "A discrete choice model for solving conflict situations between pedestrians and vehicles in shared space," *arXiv preprint arXiv:1709.09412*, 2017.

[11] R. Schönauer, M. Stubenschrott, W. Huang, C. Rudloff, and M. Fellendorf, "Modeling concepts for mixed traffic: Steps toward a microscopic simulation tool for shared space zones," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2316, pp. 114–121, 2012.

[12] A. Graves, "Generating sequences with recurrent neural networks," *arXiv preprint arXiv:1308.0850*, 2013.

[13] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social lstm: Human trajectory prediction in crowded spaces," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 961–971.

[14] X. Wang, R. Jiang, L. Li, Y. Lin, X. Zheng, and F.-Y. Wang, "Capturing car-following behaviors by deep learning," *IEEE Transactions on Intelligent Transportation Systems*, 2017.

[15] K. Yamaguchi, A. C. Berg, L. E. Ortiz, and T. L. Berg, "Who are you with and where are you going?" in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1345–1352.

[16] P. Trautman, J. Ma, R. M. Murray, and A. Krause, "Robot navigation in dense human crowds: the case for cooperation," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 2153–2160.

[17] M. Gérin-Lajoie, C. L. Richards, and B. J. McFadyen, "The negotiation of stationary and moving obstructions during walking: anticipatory locomotor adaptations and preservation of personal space," *Motor control*, vol. 9, no. 3, pp. 242–269, 2005.

[18] M. Wirz and D. Roggen, "Detecting pedestrian flocks by fusion of multi-modal sensors in mobile phones," in *ACM Conference on Ubiquitous Computing*, 2012, pp. 240–249.

[19] H. Cheng and M. Sester, "Mixed traffic trajectory prediction using lstm–based models in shared space," in *The Annual International Conference on Geographic Information Science*. Springer, 2018, pp. 309–325.

[20] S. Hacohen, N. Shvalb, and S. Shoval, "Dynamic model for pedestrian crossing in congested traffic based on probabilistic navigation function," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 78–96, 2018.

[21] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz, "The walking behaviour of pedestrian social groups and its impact on crowd dynamics," *PloS one*, vol. 5, no. 4, p. e10047, 2010.

[22] J. R. Munkres, *Topology*. Prentice Hall, 2000.

[23] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.