


Roadside pedestrian motion prediction using Bayesian methods and particle filter

Qing Xu | Haoran Wu  | Jianqiang Wang | Hui Xiong | Jinxin Liu | Keqiang Li

School of Vehicle and Mobility, Tsinghua University,
Beijing, People's Republic of China

Correspondence

Jianqiang Wang, School of Vehicle and Mobility,
Tsinghua University, Beijing 100084, People's Republic of China.

Email: wjqlws@tsinghua.edu.cn

Funding information

National Natural Science Foundation of China,
Grant/Award Numbers: 52072212, 52072214;
National Science Fund for Distinguished Young
Scholars, Grant/Award Number: 51625503;
Tsinghua-Alliance Joint Research Center for Intelligent
Mobility, Grant/Award Number: 20193910045

Abstract

Accidents between vehicles and pedestrians account for a large partition of severe traffic accidents. So, pedestrian motion prediction becomes a major concern for intelligent vehicles. However, current researches often neglect pedestrian behaviour and/or intention in motion prediction. Meanwhile, related works are scattered and divided into many small fields. No integrated system is proposed to connect the task of perception and decision. To solve these problems, a pedestrian motion prediction model is proposed in this paper. The proposed method predicts pedestrian motion based on the combination of pedestrian crossing behaviour and intention. Pedestrian behaviour is recognized using the Bayesian posterior model, and pedestrian intention is recognized by the dynamic Bayesian network. A modified particle filter and a behavioural motion model are used to integrate the behaviour and intention into motion prediction. The effectiveness of the proposed method is verified in our provided BPI dataset with eight typical scenarios defined by road type, vehicle velocity etc. The results show that this method can give an accurate distribution of pedestrians' future trajectories.

1 | INTRODUCTION

Pedestrians are the most vulnerable road users in traffic accidents. In 2018, an Uber intelligent vehicle caused the death of a pedestrian in their road test in Arizona, which became the world's first pedestrian fatality accident caused by intelligent vehicles [1]. Hence, pedestrian protection systems attract more and more attention from researchers and manufacturers in recent years. However, current pedestrian protection systems mainly focus on pedestrian detection and early warning for immediate collisions [2], which makes false braking and insufficient braking easy to occur. To help intelligent vehicles make an easier and better choice, the assistant systems should predict pedestrian motion before fatal situations would occur.

According to the research goals, the prediction of pedestrian motion can be divided into pedestrian behaviour/intention recognition and pedestrian trajectory prediction. Here, intention refers to the Crossing/Not-Crossing (C/NC) decision of the

pedestrian. The behaviour/intention recognition concentrates on the global states of pedestrians, while the trajectory prediction concerns the future distributions of pedestrian velocity and position. Based on the prediction horizon, the trajectory prediction methods can be divided into short-term trajectory prediction and long-term trajectory prediction. The boundary is always taken as 1.5 s.

The division of pedestrian behaviour/intention recognition methods mainly depends on the models constructed for pedestrian motion, which include macroscopic models, microscopic models, Markov chain models etc. The macroscopic models are always used in traffic flow studies, like [3–5]. These models neglect the differences between pedestrians, which makes them difficult to be applied to pedestrian crossing scenarios.

Microscopic models study each pedestrian individually. But they are always short of experiments and are difficult to be applied to real scenarios. Gipps et al. [6] proposed a benefit-cost cellular model. Each pedestrian occupied a cell in the plane

and moved to the most beneficial cell according to the scores concerning other pedestrians. Matsushita and Okazaki et al. [7, 8] proposed a magnetic force model to introduce the concept of magnetic force into the prediction of pedestrian motion. Pedestrians and obstacles were defined as positive nodes, and the destination was defined as negative. Helbing et al. [9–11] proposed a social force model that also explained pedestrian motion in the view of force and Zeng et al. [12, 13] employed and adjusted the theory for pedestrian behaviour analysis at signalized intersections. Zhu et al. [14] and Cao et al. [15] used the social force model to study the flow characteristics or obtain an available and optimal route, but these studies concentrated on pedestrian flow rather than a single pedestrian.

The Markov chain models transform pedestrian motion into the Markov process and discretize pedestrian states into finite variables. The transitions between different variables are in research. Wakim et al. [16, 17] simplified pedestrian motion into standing, walking, jogging, and running. The velocity distribution of each behaviour was given based on the truncated Gaussian distribution. The step size and transitions between different states were also defined. Compared with microscopic models, the Markov chain models are easier to explain and apply to real scenarios. Besides, cellular automaton model [18], fuzzy logic models [19, 20], machine learning models [21–23, 42], dynamic Bayesian network (DBN) [24], expectation-maximization algorithm [25] were also used in pedestrian behaviour/intention recognition.

Short-term trajectory prediction regards pedestrian motion as a continuous process relating to the previous motion, and builds models to simulate pedestrian movement. Conventional rigid motion models include constant speed, constant acceleration, constant velocity, and constant turn. Based on the above models, Kalman filter-based methods are adopted to track and predict pedestrian position [26]. For linear pedestrian motion models, the Kalman filter can precisely predict pedestrian motion. For nonlinear pedestrian motion models, extended Kalman filter and unscented Kalman filter perform better. To overcome the insufficient prediction accuracy with a single pedestrian motion model, interacting multiple model Kalman filter is widely used to combine different motion models for prediction. However, different from rigid motion models, pedestrian motion is more arbitrary [27]. Motion models considering pedestrian kinematic characteristics attract more attention in these years. Goldhammer et al. [28] studied pedestrian's starting process at intersections with traffic lights. With the analysis of collected data, pedestrian minimum steady movement speed was defined and pedestrian starting was described with a sigmoid model and a subsection model. Iryo-Asano et al. [29] considered the influence of traffic lights on pedestrian's velocity at sidewalks. The difference between crossing speed and current speed, along with the flash of green light, had a significant influence on pedestrian speed.

The above models are all first-order Markov models, and they mainly take pedestrian positions as the states. To describe pedestrian movement more accurately, high-order and nonlinear Markov models are proposed integrating images and pedestrian postures. Keller et al. [30] proposed a first-order non-

linear pedestrian motion model based on Gaussian process dynamic models. Principle components analysis and machine learning were used to get the model parameters. They also proposed a high-order nonlinear pedestrian motion model based on probabilistic hierarchical trajectory matching [31]. This model extracted pedestrian movement characteristics from the depth disparity map and the optical flow map. Then they combined these characteristics with the horizontal and vertical positions of pedestrians to give the trajectory prediction results. Quintero et al. [32] built a pedestrian motion model based on the Gaussian process dynamic model to reduce the dimension of pedestrian 3d attitude. They also improved the method by reducing the 3d time-related information to key points and joints to divide pedestrian motion into four types of activities [41]. The best-suited model will be selected as the motion model for the following trajectory prediction. Besides, some researches also concerned the context in short-term trajectory prediction. Kooij et al. [33] combined the dynamic Bayesian network with a switching linear dynamic system to obtain pedestrian future positions with the context. Pedestrian crossing intention was recognized before the prediction.

Long-term trajectory prediction is not limited to the future positions of pedestrians. It focuses more on pedestrian route choice in a relatively long time. So, it's less relative to the previous trajectory, and more relying on pedestrian intention and the situation. Rehder et al. [34] treated pedestrian motions as latent variables, then transformed the motion prediction problem into a trajectory planning work. The method could be adaptive to different motion models and different environments. Karasev et al. [35] included the destination of the pedestrian in the proposed motion model. The orientation of the pedestrian changed with the destination while the velocity of the pedestrian changed with the motion model. Bandyopadhyay et al. [36] used a partially observable Markov decision process to model pedestrian latent destination distribution. Then, traditional motion models were used to predict the trajectory. Sarah et al. [37] proposed a switching point detection and clustering method based on the traditional Gaussian process model. Unsupervised learning was combined with the Gaussian process to quickly detect the changes in pedestrian intention. Rasouli et al. [43] proposed an RNN encoder-decoder architecture for pedestrian trajectory prediction. Pedestrian identified latent intention and the predicted vehicle speed were combined to the trajectory prediction module by the decoder. In conclusion, long-term pedestrian trajectory prediction needs a strong perception to get the pedestrian position, velocity, postures, and the environment. Limited by the accuracy of sensors, real-world applications are still immature.

In this paper, we focus on roadside pedestrian motion prediction at vehicle-pedestrian interactive scenarios without traffic lights and crosswalks, aiming at building an integrated system combining pedestrian behaviour recognition, intention recognition, and trajectory prediction together. To address this issue, we put forward a novel method in which pedestrian behaviour, intention, and context are recognized then integrated into the motion prediction to get better prediction results. We demonstrate how to consider pedestrian characteristics and

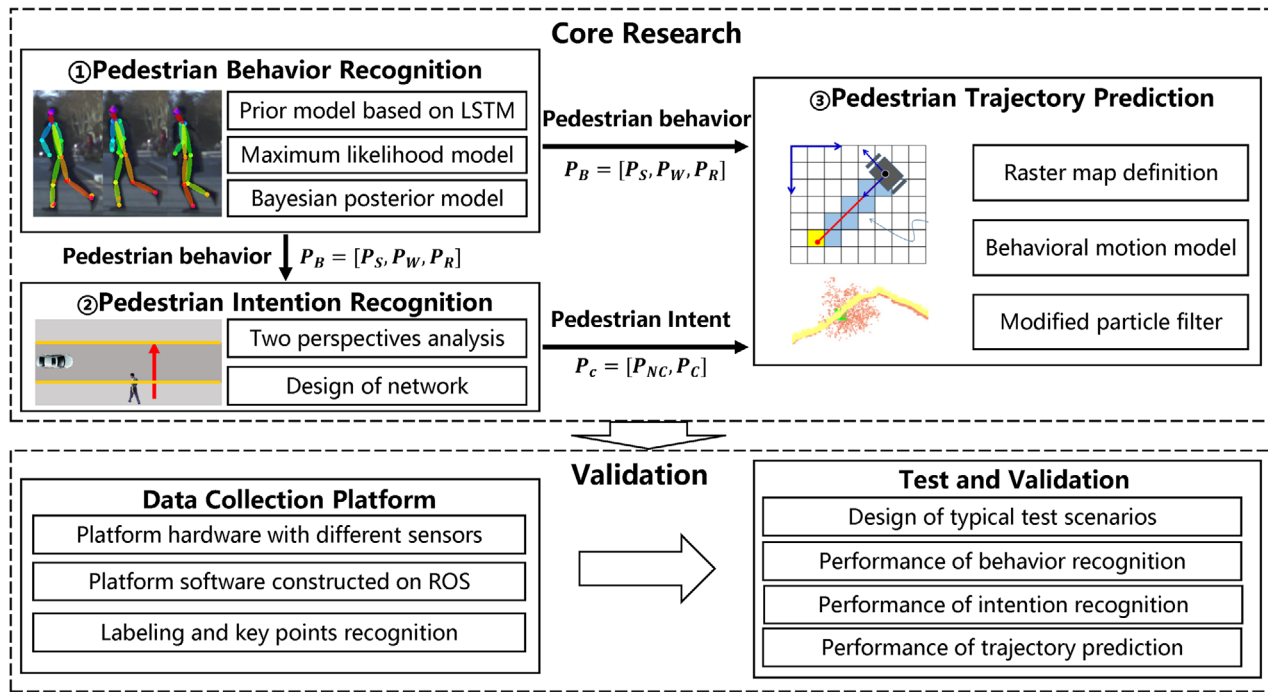


FIGURE 1 The framework of our progress. Pedestrian behaviour will be used as the input of intention recognition. Then particle filter will be used to combine the behaviour and intention to get the trajectory prediction results. A data collection Platform is built for method validation in our provided dataset

environment features concurrently in trajectory prediction. Firstly, Bayesian methods and neural networks are used in pedestrian behaviour recognition and intention recognition to improve recognition performances. Then, the recognized behaviour and intention are integrated into the modified particle filter to get better trajectory prediction results. Specifically, our contributions are as follows:

1. The notion of the integrated method is proposed for roadside pedestrian motion prediction at vehicle-pedestrian interactive scenarios without traffic lights and crosswalks. Rather than recognizing pedestrian intention and predicting its trajectory respectively, we provide an explicit method to combine pedestrian behaviour, intention, and context in the motion prediction. Our experiments reveal that the integrated method could get rational motion prediction results in typical scenarios.
2. For pedestrian behaviour recognition, the Bayesian posterior model is proposed to mix the merits of the prior behaviour model based on long short-term memory (LSTM) and the maximum likelihood behaviour model based on velocity distribution. Our posterior model aims at fully utilizing both the information of pedestrian pose and pedestrian motion to eliminate the influence of occlusion and improve the behaviour recognition accuracy.
3. For pedestrian intention recognition, the DBN based on the Markov assumption is proposed to recognize pedestrian crossing intention, which regards the recognized behaviours as an observed variable. The network construction derives from the analysis both from the perspective of pedestrians

and the perspective of vehicles, which makes it suitable for vehicle-pedestrian interactive scenarios.

The remaining of this paper is organized as follows. Section 2 introduces the framework of the proposed method and defines some important terms. The methodology of this research is presented in Section 3. Pedestrian behaviour recognition, intention recognition and trajectory prediction are introduced in Sections 3.1–3.3 respectively. The foundation of the experiment platform and the verification results are shown in Section 4. Preparation for data collection and design of test scenarios are presented in Sections 4.1–4.2. And the verification results are listed in Sections 4.3–4.5. Finally, Section 5 concludes this paper.

2 | RESEARCH FRAMEWORK

As mentioned above, this research focuses on roadside pedestrian motion prediction, and the framework of our progress is shown in Figure 1. In this research, “recognition” is based on collected historical information to recognize pedestrian current motion and its C/NC decision, and “prediction” is based on the recognition of the current state to predict where the pedestrian position will be in the future. To clearly distinguish pedestrian physical movement from spiritual thinking, we define “behaviour” as a pedestrian physical characteristic, like standing, walking, and running. And “intention” is defined as the decision of the pedestrian on whether or not to cross the road (C/NC problem). Besides the core research, we also build a platform for data collection and design typical test scenarios to validate

the performances of the algorithms. The contents of the core research are as follows:

1. Pedestrian behaviour recognition is designed to recognize pedestrian standing, walking, and running behaviour and output its probability distribution $[P_S, P_W, P_R]$. A prior model using pedestrian pose info and a maximum likelihood model using pedestrian motion info are built based on LSTM and data fitting method respectively, then they are combined using the Bayesian method to get better results and eliminate the influence of occlusion.
2. Pedestrian crossing intention recognition is designed to recognize pedestrian crossing intention and output the probabilities of not-crossing and crossing $[P_{NC}, P_C]$ in the whole process. In this part, a dynamic Bayesian network is designed considering different perspectives of intelligent vehicle and pedestrian itself to make it suitable for vehicle-pedestrian interactive scenarios.
3. Pedestrian trajectory prediction considering the probability distribution of pedestrian behaviour, pedestrian crossing intention, and the context is designed based on a modified particle filter to output pedestrian trajectory with the distribution of particles. The importance sampling of particles is changed based on the attributes of the environment.

3 | METHODOLOGY

3.1 | Pedestrian behaviour recognition

Pedestrian behaviour has various presentations in the traffic environment. Considering the major difference in pedestrian pose and velocity, it is classified into standing, walking, and running in this research. Thus, we have:

$$B = \{S, W, R\} \quad (1)$$

where B is the discrete set of pedestrian behaviour. S , W , and R are standing, walking, and running, respectively.

As we have claimed, pedestrian behaviour can be classified both by pedestrian pose and motion. Pose-based classification takes advantage of images and reacts faster to the change of pedestrian behaviour. Besides, the classification is less affected by vehicle motion, since the pedestrian pose from images keeps relatively stable while the ego vehicle moves. However, this method highly depends on the image quality and observation angle of pedestrians. When the target pedestrian is blocked, the difficulty of classification shall rise a lot. Motion-based classification is much straighter since there is an explicit relationship between pedestrian motion (especially velocity) and our defined behaviour. The fact that pedestrian velocity can be derived from various sensors also improves the robustness of this method. Unfortunately, in most cases, pedestrian velocity is obtained via tracking or other methods other than direct sensing, which leads to low precision and large fluctuation. Compared to pose-based classification, it's also affected by the motion of ego vehicle. The vehicle velocity is generally much higher than pedestrian, and

pedestrian velocity calculated from the vehicle coordinate system will deviate from the expectation.

To overcome the disadvantages of these two methods, we build a prior model using pedestrian pose as input and a maximum likelihood model using pedestrian velocity as input. The prior model outputs a probabilistic estimation of pedestrian behaviour as the prior probability $\tilde{P}(B)$, and the maximum likelihood model contains a conditional probability $P(V|B)$ between pedestrian behavior and velocity, then the maximum likelihood probability $\hat{P}(B)$ under a certain velocity is given. Finally, the Bayesian method enables us to calculate the posterior probability $\hat{P}(B)$ as Equation (2):

$$\hat{P}(B) \propto P(V|B)\tilde{P}(B) \quad (2)$$

Let $P(B)$ be the final probability distribution of the behavior, it's calculated with Equation (3):

$$P(B) = \begin{cases} \hat{P}(B) \tilde{P}(B), & \tilde{P}(B) \text{ both available} \\ \hat{P}(B) & \text{only } \tilde{P}(B) \text{ available} \\ \tilde{P}(B) & \text{only } \hat{P}(B) \text{ available} \end{cases} \quad (3)$$

When the prior probability $\tilde{P}(B)$ and maximum likelihood probability $\hat{P}(B)$ are both available, the posterior probability $\hat{P}(B)$ is set to $P(B)$. Otherwise, the available one will be set to $P(B)$ directly. This design will expand the working condition of the behavior recognition algorithm and improve the prediction performance.

3.2 | Prior model based on long short-term memory

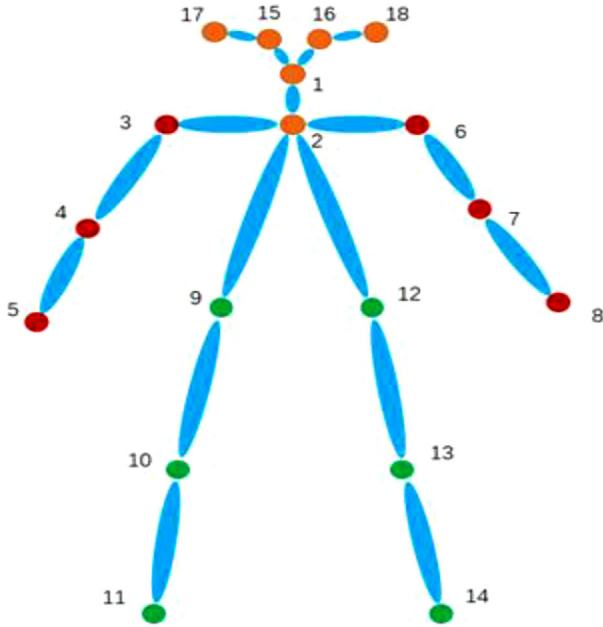
Pedestrian pose is used as the input of the prior model. Traditional image-based pedestrian behaviour recognition methods use a single frame original image as input and learn pedestrian behaviour by supervised learning. However, two negative aspects of these methods must be considered: First, training original image requires a large model and massive data, otherwise the effect cannot be guaranteed; second, a single frame image is not able to show pedestrian behaviour precisely. Pedestrian motion is a continuous process, different behaviours may seem similar in a single frame. To reduce computation cost and model complexity, and take the continuous process of pedestrian behaviour into consideration, this research uses pedestrian 18 key points as input and the LSTM to recognize pedestrian behaviour [38], the 18 key points are shown in Figure 2, and the definition of these points can be seen in Table 1. Since the LSTM model is widely known, only model input will be introduced in detail.

$$x = [p_x^1, p_y^1, p_x^2, p_y^2, \dots, p_x^{18}, p_y^{18}] \quad (4)$$

The input vector is composed of longitudinal and lateral positions (p_x^i, p_y^i) of the 18 defined key points for the studied

TABLE 1 Index of pedestrian key points

	1	2	3	4	5	6	7	8	9
Index	10	11	12	13	14	15	16	17	18
Name	Mouth	Neck	Left shoulder	Left elbow	Left wrist	Right shoulder	Right elbow	Right wrist	Left hip
	Left knee	Left ankle	Right hip	Right knee	Right ankle	Left eye	Right eye	Left ear	Right ear

**FIGURE 2** Pedestrian key points

pedestrian, as shown in Equation (4). The superscripts represent the key points indices and the subscripts mean the coordinate axes.

Influenced by the pedestrian height and image acquisition range, the coordinates of the key points differ in different frames, which strongly influences the consistency of the input vectors. To improve this problem, the key points in different frames are normalized by the length of the pedestrian's torso. The coordinates of the upper and lower points are calculated in Equations (5) and (6):

$$p_i^{upper} = \frac{p_i^{l_shoulder} + p_i^{r_shoulder}}{2}, \quad i = x, y \quad (5)$$

$$p_i^{lower} = \frac{p_i^{l_hip} + p_i^{r_hip}}{2}, \quad i = x, y \quad (6)$$

where $(p_i^{l_shoulder}, p_i^{r_shoulder})$, $(p_i^{l_hip}, p_i^{r_hip})$ represent the coordinates of pedestrian left and right shoulder, left and right hip, respectively. And the length of pedestrian's torso is calculated as:

$$\Delta d = \sqrt{\left(p_x^{upper} - p_x^{lower}\right)^2 + \left(p_y^{upper} - p_y^{lower}\right)^2} \quad (7)$$

The 18 key points are normalized by the length of pedestrian's torso and the lower point is used as the central point. The normalised position $(\hat{p}_x^i, \hat{p}_y^i)$ and the normalized input vector \hat{x} are shown in:

$$\begin{aligned} \hat{p}_x^i &= \frac{p_x^i - p_x^{lower}}{\Delta d}, \quad i = 1, 2, \dots, 18 \\ \hat{p}_y^i &= \frac{p_y^i - p_y^{lower}}{\Delta d}, \quad i = 1, 2, \dots, 18 \\ \hat{x} &= [p_x^1 \ p_y^1 \ p_x^2 \ p_y^2 \ \dots \ p_x^{18} \ p_y^{18}] \end{aligned} \quad (8)$$

3.3 | Maximum likelihood model based on data fitting method

Different from the LSTM method, the maximum likelihood model constructs a function between pedestrian behaviour and velocity distribution to complete the behaviour recognition. Without consideration of pedestrian orientation, pedestrian behaviour is identified through different velocity distribution from empirical rules and captured data.

3.3.1 | Standing

Theoretically, the velocity of the standing pedestrian is zero. However, in real tests, the measured speed is always influenced by noises and system errors. So, in this research, the velocity is explained with an exponential distribution:

$$P(v|B = S) = \begin{cases} \lambda_s e^{-\lambda_s v} & v \geq 0 \\ 0 & v < 0 \end{cases} \quad (9)$$

where $P(v = 0|B = S) = \lambda_s > 0$, and the function is strictly decreasing.

3.3.2 | Walking

When a pedestrian is walking, the velocity always remains stable with some deviation from the average speed. So, we choose the Gaussian distribution to explain the speed distribution:

$$P(v|B = S) = \frac{1}{\sqrt{2\pi}\sigma_w} \exp\left(-\frac{(v - \mu_w)^2}{2\sigma_w^2}\right) \quad (10)$$

where μ_v means the expectation of the average speed and σ_v represents the standard deviation.

3.3.3 | Running

When the pedestrian is running, the velocity distribution will change at high speed. When pedestrian speed exceeds the general range, the probability density will quickly decrease to zero. So, we use Weibull distribution to describe it:

$$P(v|B = S) = \frac{k_r}{\lambda_r} \left(\frac{v}{\lambda_r}\right)^{k_r-1} \exp\left(-\left(\frac{v}{\lambda_r}\right)^{k_r}\right) \quad (11)$$

where k_r , λ_r are shape and scale parameters respectively.

Based on the maximum likelihood model, we can calculate the likelihood of each behaviour with Equation 12), then the posterior probability $\hat{P}(B)$ can be calculated with Equation 2 to get the final distribution of each behaviour.

$$\hat{P}(B = \%i) = \frac{P(v|B = i)}{\sum_k P(v|B = k)} \% \quad i, k \in \{S, \%W, \%R\} \quad (12)$$

3.4 | Pedestrian intention recognition

Pedestrian crossing intention is influenced by many factors. On the one hand, pedestrian crossing depends on its observation of road structure, obstacles, vehicles, and other traffic participants. Pedestrian demand also determines its destination. On the other hand, vehicles cannot sense the intention of pedestrians. Vehicles can only infer pedestrian motion from pedestrian posture, context etc. In this part, we will analyse the possible factors that may influence pedestrian motion from pedestrian and vehicle perspectives, respectively.

3.4.1 | Pedestrian perspective

In typical pedestrian crossing scenarios (with vehicle passing), there are three factors that will help justify pedestrian crossing intention: pedestrian destination, scene criticality, and whether the pedestrian is in the crossing area. Pedestrian destination is the main factor affecting pedestrian crossing intention. The pedestrian will not cross if its destination is not on the opposite side. Scene criticality means the danger pedestrian will face when trying to cross. It mainly depends on the relative distance and velocity between pedestrian and other traffic participants. "Crossing area" represents the areas where pedestrians are regarded as crossing when they get in this area. In this research, it refers to the roadway. Whether the pedestrian is in the crossing area is the external performance of pedestrian crossing behaviour. Pedestrians in the crossing area are more likely to complete their crossing process rather than step back to the roadside. These three factors constitute the foundation of the explanation model of pedestrian intention jointly.

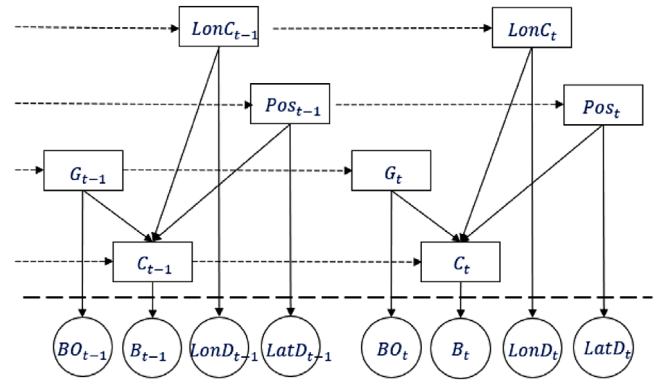


FIGURE 3 Dynamic Bayesian network for pedestrian intention recognition [33]

3.4.2 | Vehicle perspective

As mentioned above, vehicles can not perceive pedestrian crossing intention directly. The observable factors that can help justify pedestrian intention include:

- Pedestrian behaviour: the posterior probability $\hat{P}(B)$ of standing, walking, and running, which is obtained in III-B
- Passable areas: areas that pedestrians can be on, like sidewalks. Pedestrian destination is generally in passable areas
- Longitudinal relative distance: the longitudinal distance between pedestrian and passing vehicle
- Lateral relative distance: the lateral distance between pedestrian and passing vehicle
- Vehicle heading angle: heading angle of ego vehicle, generally zero
- Vehicle speed: the speed of ego vehicle
- Pedestrian heading angle: heading angle of the pedestrian
- Pedestrian speed: the speed of the pedestrian
- Time to Collision: when the pedestrian and vehicle will be in collision if they remain at the current speed

These factors constitute the foundation of the description model of pedestrian intention jointly.

Based on the above analysis, we take the factors from pedestrian and vehicle perspectives as hidden variables and observed variables respectively to construct the dynamic Bayesian network for pedestrian intention recognition, as shown in Figure 3.

The rectangle nodes mean hidden variables and the circle nodes mean observed variables. The meaning of all these parameters is shown in Table 2. To facilitate the introduction of the following process, the condition variables set E and the observation variables set O are introduced in Equation 13).

$$\begin{aligned} E &= \{G, LonC, Pos\} \\ O &= \{BO, LonD, LatD\} \end{aligned} \quad (13)$$

TABLE 2 Parameters of the dynamic Bayesian network

Variables		Meaning	Domains
Hidden variables	G	Opposite destination	[True, False]
	LonC	Longitudinal danger exists	[True, False]
	Pos	Pedestrian in crossing area	[True, False]
	C	Pedestrian crossing intention	[True, False]
Observed variables	BO	Body orientation	[0, 360]°
	LonD	Pedestrian longitudinal distance to vehicle	[0, 40]m
	LatD	Pedestrian distance to road curb	[-10, 10]m
	B	Pedestrian behaviour	[S, W, R]

where the set E includes all hidden variables except for pedestrian intention C , and the set O includes all observed variables except for pedestrian behavior B .

The inference process of DBN applied in crossing intention recognition can be divided into two stages: predict and update [33]. Two kinds of probability are introduced in the inference process, the prior probability $\bar{P}(\cdot)$ and the posterior probability $\hat{P}(\cdot)$. The inference process depends on the Markov assumption. In the prediction stage at time t , the prior probability $\bar{P}(C_t, E_t)$ is calculated based on the posterior probability $\hat{P}(C_{t-1}, E_{t-1})$ at time $t-1$, as shown in Equation (14):

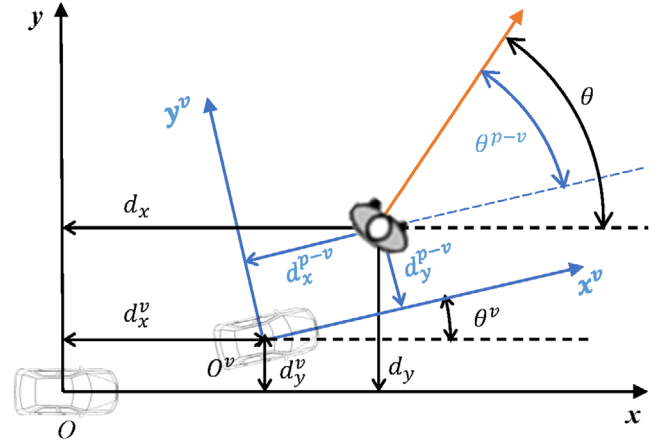
$$\begin{aligned} \bar{P}(C_t, C_{t-1}, E_t, E_{t-1}) &= P(C_t | C_{t-1}, E_t) \times P(E_t | E_{t-1}) \\ &\quad \times \hat{P}(C_{t-1}, E_{t-1}) \\ \bar{P}(C_t, E_t) &= \sum_{C_{t-1}} \sum_{E_{t-1}} \bar{P}(C_t, C_{t-1}, E_t, E_{t-1}) \end{aligned} \quad (14)$$

In the update stage at time t , the main work is to obtain the posterior probability $\hat{P}(C_t, E_t)$ depending concurrently on the observations at time t and the results at time $t-1$. Considering the results at time $t-1$ have been processed in the prediction stage, we can get Equation (15) by Bayesian formula:

$$\begin{aligned} \hat{P}(C_t, E_t) &\propto P(O_t | E_t) \times P(B_t | C_t) \times \bar{P}(C_t, E_t) \\ \hat{P}(C_t) &= \sum_{E_t} \hat{P}(C_t, E_t) \end{aligned} \quad (15)$$

3.5 | Pedestrian trajectory prediction

The pedestrian motion model is the basis of pedestrian trajectory prediction. Tradition motion models treat the pedestrian motion as rigid body motion and restrict it to x - y coordinates. However, pedestrian speed always changes in the direction of pedestrian body orientation, along with some turning. Pedestrian behaviour will also influence pedestrian motion. According to the analysis, we employ the motion model proposed by

**FIGURE 4** Relative position of pedestrian and vehicle

Hashimoto et al. [39]:

$$\begin{aligned} \begin{bmatrix} d_x \\ d_y \end{bmatrix}_k &= \begin{bmatrix} d_x \\ d_y \end{bmatrix}_{k-1} + v_{k-1} \Delta t \begin{bmatrix} \cos \theta_{k-1} \\ \sin \theta_{k-1} \end{bmatrix} \\ \theta_k &= \theta_{k-1} + \delta_\theta \\ v_k &= \begin{cases} |0 + \delta_v| & \text{if } P(B = S) > 0.5 \\ |v_{k-1} + \delta_v| & \text{others} \end{cases} \end{aligned} \quad (16)$$

where v_{k-1} and θ_{k-1} mean pedestrian velocity and orientation at the prediction step $k-1$, $(d_x, d_y)_k$ represent pedestrian position at the prediction step k . The equations are constructed on the assumption that pedestrian velocity only changes in the direction of pedestrian orientation. δ_θ and δ_v are the additive noises on pedestrian orientation and velocity. When pedestrian behavior is standing, pedestrian velocity is set to zero with δ_v .

The employed motion model can clearly explain pedestrian motion in the global coordinate system. However, in this research, the observation data are collected through vehicle-mounted cameras and LIDAR, it's essential for us to transform pedestrian motion to the vehicular coordinate system. The transformation relations from the global coordinate system to the vehicular coordinate system are shown in Figure 4.

where $x-O-y$ is the global coordinate system and $x^v-O^v-y^v$ is the vehicular coordinate system, (d_x^v, d_y^v, θ^v) mean vehicle position and heading angle in the global coordinate system. And $(d_x^{p-v}, d_y^{p-v}, \theta^{p-v})$ represent pedestrian position and orientation in the vehicular coordinate system. The pedestrian observation model from the vehicle perspective is shown in Equation 17:

$$\begin{aligned} \begin{bmatrix} d_x^{p-v} \\ d_y^{p-v} \end{bmatrix}_k &= \begin{bmatrix} \cos \theta_k^v & \sin \theta_k^v \\ -\sin \theta_k^v & \cos \theta_k^v \end{bmatrix} \begin{bmatrix} d_x - d_x^v \\ d_y - d_y^v \end{bmatrix}_k + \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix} \\ \theta_k^{p-v} &= \theta_k - \theta_k^v + \delta_\theta^{p-v} \end{aligned} \quad (17)$$

TABLE 3 Basic environment features

Basic factors	Description	Index
Pedestrian	Objective single pedestrian	M1
Vehicle	Vehicle that may affect pedestrian motion	M2
Vehicle accessible areas	Vehicle accessible areas, motorway in this research	M3
Pedestrian accessible areas	Pedestrian accessible but vehicle non-accessible areas	M4
Non-crossable Areas	Areas non-crossable for all traffic participants	M5

where δ_x , δ_y and δ_{θ}^{p-v} are observation errors of the longitudinal distance, lateral distance, and pedestrian body orientation, respectively.

After defining the pedestrian motion model, scenario reconstruction is conducted. Pedestrian motion prediction depends on the description of the scenario. In this section, the raster map is used to abstract the traffic environment features. The basic environment features are shown in Table 3.

The raster map is constructed with the vehicle's initial position as the original point, size of 50 m \times 20 m. The size of each raster is 0.2 m \times 0.2 m. In the following motion prediction steps, the pedestrian trajectory is composed of grids occupied in each timestamp.

The importance sampling process of the particle filter also depends on the attribute of the grids. In the traditional particle filter, Gaussian function etc. are used to weight the difference of particles in each prediction step. However, the pedestrian future trajectory is influenced by pedestrian behaviour and intention, the distribution is not always Gaussian. In this research, the weight matrix $w(p_i|m_i, c_i)$ is defined depending on the context and pedestrian crossing intention. p_i means the weight of the i_{th} particle. $m_i \in \{M2, M3, M4, M5\}$ represents the attribute of the i_{th} particle, obtained from the position attribute of the particle (d_x, d_y) in the raster map. $c_i \in \{0, 1\}$ means the crossing intention of the i_{th} particle, which is decided by DBN. The weight function is defined in Equation 18:

$$w(p_i|m_i, c_i) = \begin{cases} 1 & \text{if } m_i = M4 \\ 0 & \text{if } m_i \in \{M2, M5\} \\ 1 & \text{if } m_i = M3 \text{ and } c_i = 1 \\ 0 & \text{if } m_i = M3 \text{ and } c_i = 0 \end{cases} \quad (18)$$

For pedestrian-accessible areas, set the weight to 1. For non-crossable areas, set the weight to 0. To vehicle accessible areas, if the pedestrian crossing intention is true, set the weight to 1, otherwise set it to 0. After calculation of all the weights, the normalization process is conducted to get the normalized weight matrix $\tilde{w}(p_i|m_i, c_i)$, as shown in Equation (19):

$$\tilde{w}(p_i|m_i, c_i) = \frac{w(p_i|m_i, c_i)}{\sum_i w(p_i|m_i, c_i)} \quad (19)$$

The basic process of particle filter [46] includes five steps: Initialization, Particle sets generation, Importance Sampling,

Resampling, Output. The prediction step S is set to 6 with 0.5s interval, and the particle number $N = 100$. The pseudo-code of the proposed algorithm is shown in Algorithm 1.

Algorithm 1 Pedestrian Trajectory Prediction Algorithm

Input: $B_{1 \times 3}, C_{1 \times 2}, STATE_{1 \times 5}, \delta_v, \delta_{\theta}$

Output: $TRAJ_{S \times N \times 5}$

$STATE_{1 \times 5} := [x_t \ y_t \ v_t \ a_t \ \theta]_{1 \times 5}$

Weight matrix $W_{1 \times N} := [1 \ 1 \ \dots \ 1]_{1 \times N}$

$STATE_0 := STATE_{1 \times 5}; W_0 := W_{1 \times N}$

for all $i = 1 : 1 : S$ **do**

$v_i = |v_{i-1} + \delta_v|; \theta_i = \theta_{i-1} + \delta_{\theta}$

if $C_{noncross} > C_{cross}$ **then**

$B_{standing} \leftarrow B_{walking}/B_{running}$ partly

if $B_{max} = B_{standing}$ **then**

$v_i = |0 + \delta_v|$

end if

end if

$TRAJ_{pre,i} \leftarrow Predictor(STATE_{i-1})$

$W_i \leftarrow Resampling(W_{i-1})$

$TRAJ_{est,i} \leftarrow Updater(TRAJ_{pre,i})$

$TRAJ_{i \times N \times 5} = TRAJ_{est,i}$

end for

The inputs of the algorithm are pedestrian behaviour probabilities ($B_{1 \times 3}$), pedestrian crossing intentions ($C_{1 \times 2}$), pedestrian states ($STATE_{1 \times 5}$), and the additive noises on pedestrian orientation and speed ($\delta_{\theta}, \delta_v$). First, the noises are added to pedestrian speed and orientation. Then, the speed is checked to see if it exceeds the maximum threshold. If so, the speed estimation is doubted and the average speed is used for trajectory prediction. Next, for each pedestrian, if its crossing intention is lower than not-crossing, the particles' walking and running behaviour will be partly changed to standing in the predictions. Then, the behaviour and previous states determine the predicted speed concurrently. Standing behaviour means no speed. Finally, a uniform particle filter is used depending on the determined speed, including the Predictor, Resampling, and Updater. The weight matrix and the trajectory are updated after each prediction step.

4 | EXPERIMENT AND RESULTS

4.1 | Preparation for data collection

Abundant information on pedestrian motion and surrounding traffic is the basis of algorithm development and verification. Focusing on different assignments, various sensors are adopted by researchers. Considering cost and reliability, companies such as Daimler and Honda choose mono camera and radar to equip their cars in the market. For researchers and companies like Google, Lidar is much more popular for its high accuracy. In this research, we also use Lidar to sense pedestrian and traffic environments. Table 4 shows the sources we use and the corresponding information that is provided by those sources. The Inertial measurement unit (IMU) is a sensor for detecting and

TABLE 4 Source of data

Source	Information
Mono camera	Pedestrian pose
Lidar	Pedestrian location, traffic environment (obstacle location, road structure etc.)
CAN	Ego vehicle velocity, steering angle, etc.
GPS/IMU	Ego vehicle acceleration, heading, etc.

**FIGURE 5** Placement of hardware on the test car

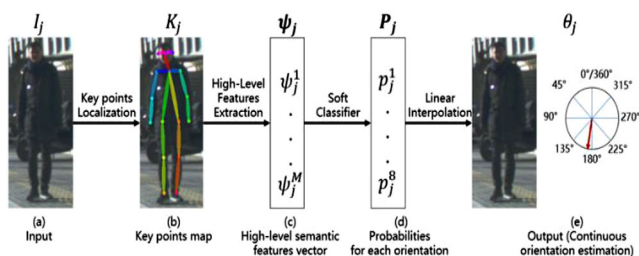
measuring the vehicle's acceleration and rotational movement. And the controller area network (CAN) is the international standardized serial communication protocol, from which we can obtain the vehicle's motion states.

The test car and the placement of the hardware are shown in Figure 5. The lidar is installed in the front of the vehicle and the mono camera is installed inside the car.

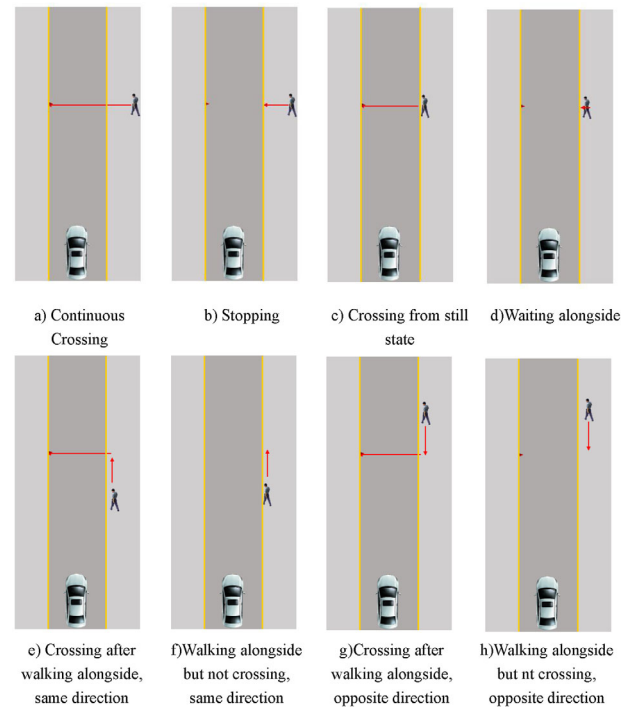
The pedestrian orientation estimation method was proposed in our previous work [40]. The principle of the estimation algorithm is shown in Figure 6, the orientation results range from 0° to 360° and they can be used for pedestrian behaviour recognition.

4.2 | Design of test scenarios

To verify the effectiveness of the proposed methods, we define 8 typical pedestrian road-crossing scenarios to collect the testing

**FIGURE 6** Pedestrian orientation estimation**TABLE 5** City road types in China

Class	Speed (km/h)	Numbers of lanes	Whole width (m)	Median
Class 1	60–90	≥ 4	40–70	must
Class 2	40–60	≥ 4	30–60	should
Class 3	30–50	2–4	20–40	can
Class 4	20–40	2	16–30	no

**FIGURE 7** Typical pedestrian crossing scenarios

data. A scenario is defined by several features, including road type, speed interval of vehicle, traffic lights, and traffic participants. In China, city roads can be divided into four classes: urban expressway, main way, support way, and subordinate way. Features of these road classes are shown in Table 5.

Restricted by road width, separation zone etc., it is possible for pedestrians to interact with vehicles only on Class 4 roads. As for traffic lights, pedestrians need to obey traffic rules because of the strong restriction. Hence, traffic lights are not considered in this research. Vehicle speed is combined with the road type. So only pedestrian motion concerns most. The defined 8 typical pedestrian crossing scenarios are shown in Figure 7. In consideration of all the mentioned features, the data collection work was conducted on an open road with a test vehicle. Volunteers observed the scene criticality and made their own decisions on whether to cross or not. The test vehicle would also change its motion by observing pedestrian behaviours.

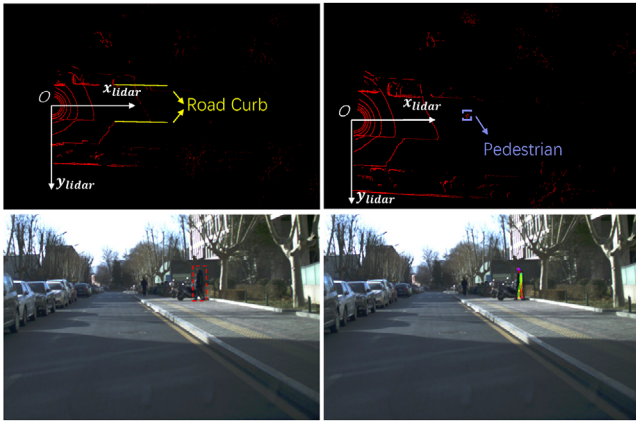


FIGURE 8 Images from the lidar and camera

4.3 | Dataset description

The BPI dataset¹ was established by Tsinghua University and was collected mainly on the campus. The vehicle's average speed ranged from 10 to 30 km/h and the road was mainly in Class 4. Pedestrians and cyclists' pose info was collected from a vehicle-mounted mono camera at 15 Hz with a resolution of 2048×1024 pixels. A lidar was used to obtain the position of the traffic participants and the environment features (like the road curb) at 10 Hz. Besides, IMU and CAN also included the vehicle information. We announced 12651 frames in 25 sequences for labelled pedestrian crossing-road scenarios from 120 captured sequences and divided them into eight typical scenarios. Cyclists' crossing intersection trajectories were also added to the scenario for possible VRU protection researches, which also included 25 sequences. The length of these sequences was between 300 frames to 800 frames. The images from the lidar and camera are presented in Figure 8.

In Figure 8, the images have been scaled and intercepted for a better exhibition. In the raw data, the resolution of the lidar picture was 4000×1000 pixels, the origin of the lidar coordinate system was at (1600, 500), and the distance between adjacent pixels was 0.05 m. The positions of traffic participants and road structures were labelled in the picture. The camera was used for pedestrian pose estimation and obtained the 18 key points of the pedestrian.

4.4 | Validation of pedestrian behaviour recognition

The verification of pedestrian behaviour recognition algorithm is constructed of the prior model based on LSTM, the maximum likelihood model based on data fitting, and the posterior model based on the Bayesian method. The LSTM model was constructed in the structure of TensorFlow, and the parameters

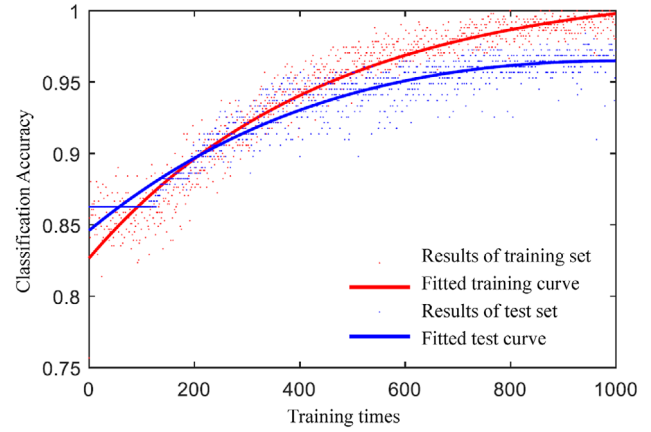


FIGURE 9 Accuracy in the training set and test set

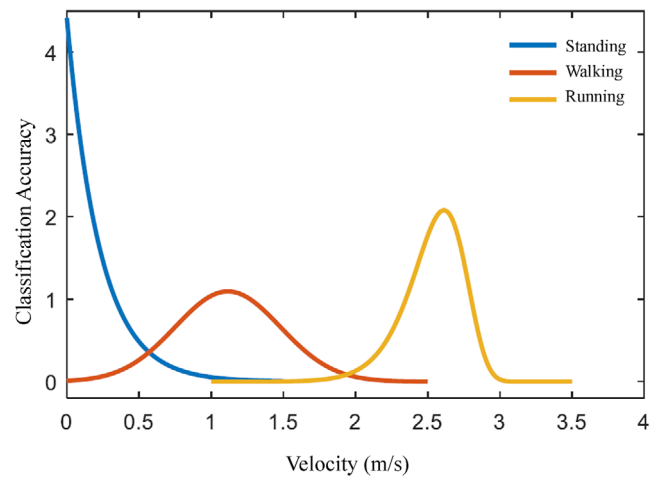


FIGURE 10 Velocity distribution of different pedestrian behaviours

were obtained using the training data or predefined. We divided the training set and the test set into 4:1. To consider both time and effect, the batch size was set to 5. In the training process, data were selected randomly as input. The gradient descent method made the learning rate decrease with the increase of training time, which helped to early overstep the local optimum and later decrease the oscillation near the extreme value. The training results were shown in Figure 9.

Along with the training, the accuracy in the training set increased to nearly 1. However, in the test set, the accuracy stopped at 0.96 after about 800 training, caused by overfitting. The best result was chosen for method comparison.

The maximum likelihood model was constructed based on pedestrian velocity. Based on the distribution of velocity, the model parameters were calibrated at 0.1 m/s intervals. The fitting curves of three pedestrian behaviours were in Figure 10. Pedestrian standing behaviour fitted the exponential distribution with $\lambda_s = 4.116$, $E(vB = S) = \sigma(vB = S) = 0.243\text{m/s}$. Pedestrian walking behavior fitted the Gaussian distribution with $\mu_w = 1.115$, $\sigma_w = 0.364$, $E(vB = W) = 1.115\text{m/s}$, and $\sigma(vB = W) = 0.364\text{m/s}$. Pedestrian running behavior fitted the Weibull distribution with $\lambda_r = 2.625$,

¹The dataset is available for non-commercial research purposes. Link: github.com/wuhaoran111/BPI_Dataset

TABLE 6 Recognition accuracy of different behaviours

Model	Pedestrian behaviour likelihood model posterior model		
	Standing	Walking	Running
LSTM model	0.97	0.97	0.97
Likelihood model	0.92	0.90	0.98
Posterior model	0.98	0.97	0.98

TABLE 7 Recognition precision of different behaviours

Model	Pedestrian behaviour likelihood model posterior model		
	Standing	Walking	Running
LSTM model	0.90	0.97	0.98
Likelihood model	0.50	0.98	0.89
Posterior model	0.86	0.98	0.98

$k_r = 14.8$, $E(vB = W) = 2.534$ m/s, and $\sigma(vB = W) = 0.210$ m/s. Pedestrian different behaviors would present obvious different distributions. So, it was effective to distinguish these behaviors from velocity distribution.

The Bayesian posterior model combined the above two behaviour models together to get the final recognition result. The recognition accuracy of these three models was shown in Table 6. The Bayesian posterior model got the best accuracy combining the prior model and the maximum likelihood model, and the LSTM model also got relatively high accuracy compared with the maximum likelihood model. However, the maximum likelihood model performed better in recognition of the running behaviour.

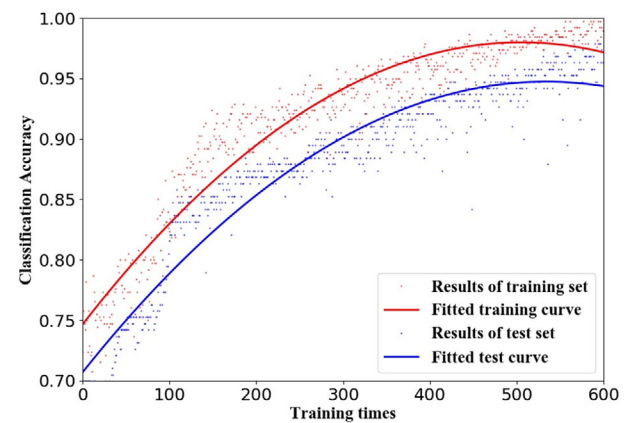
The comparison of recognition precision was shown in Table 7. Precision represented the model's credibility when giving a positive judgment of the behaviour. For each behaviour, the Bayesian posterior model performed well and the maximum likelihood model performed relatively worse. It was worth attention that the three methods' recognition precision for standing behaviour was apparently lower than other behaviours. On the one hand, it was caused by the vague boundary of standing and slowly walking. On the other hand, the scale of positive standing behaviour data was relatively small. The precision could be improved by more training and more clear classification features.

The comparison of recognition rate was shown in Table 8. The recognition rate represented the recognition ability of the model to positive samples. The maximum likelihood model got a good recognition rate of standing behaviour. Along with the previous observation, we could know that the maximum likelihood model held high tolerance to the velocity of standing, which also led to a more erroneous judgment of other behaviours.

The Bayesian posterior model got the best recognition result in most of the comparisons in the BPI dataset. To further prove

TABLE 8 Recognition rate of different behaviours

Model	Pedestrian behaviour likelihood model posterior model		
	Standing	Walking	Running
LSTM model	0.96	0.97	0.96
Likelihood model	0.92	0.91	0.82
Posterior model	0.97	0.97	0.96

**FIGURE 11** Accuracy in the PIE dataset

the adaptability of the proposed method, the behaviour recognition method was also applied to the PIE dataset [43] in the experiment. As the key points of the pedestrian were difficult to capture at low resolution, only pedestrians whose bounding boxes width was larger than 20 pixels were used for the behaviour recognition. Standing and walking were recognized according to the labelled behaviour. The results were shown in Figure 11.

At this time, after 500 epochs of training, the training curve started to become overfitting, and the accuracy of the behaviour recognition reached 93% for the LSTM network. The maximum likelihood model only obtained 80% accuracy for the standing behaviour and 79% accuracy for the walking behaviour. Combining the two models together, the Bayesian posterior model got 90% accuracy for the behaviours. The performance of the posterior model is worse than the LSTM network in this scenario. However, this does not mean that the posterior behaviour model is useless. In many scenarios, pedestrian key points may get sheltered by obstacles, then the maximum likelihood model can be used as insurance.

4.5 | Validation of pedestrian intention recognition

To show the performance of our method in different scenarios, 25 typical sequences were selected for eight defined pedestrian typical scenarios. Pedestrian intention recognition was conducted for all these sequences. To compare the results in the

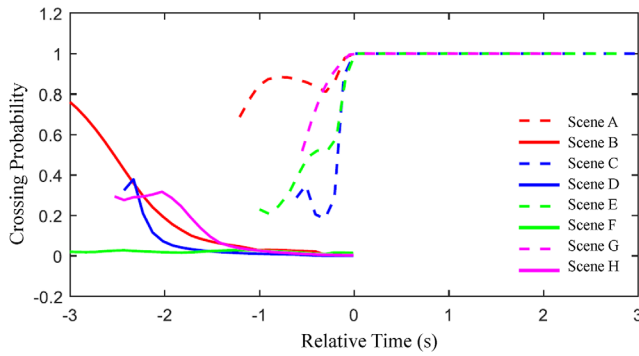


FIGURE 12 Pedestrian crossing intention recognition in eight typical scenarios

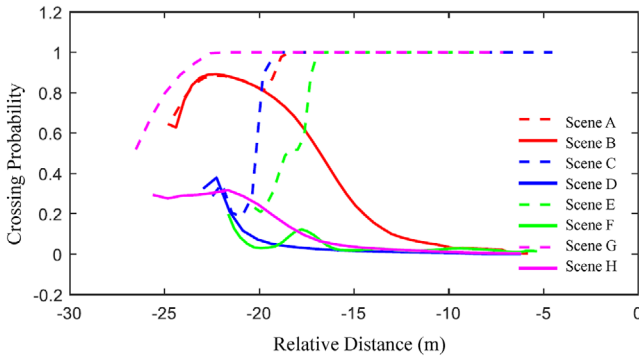


FIGURE 13 Pedestrian crossing intention recognition with relative distance

time domain, the zero points were defined as the moment of crossing in crossing scenarios. In non-crossing scenarios, the zero points were defined as the moment of the final frame. Pedestrian crossing intention recognition results were shown in Figure 12.

In all pedestrian crossing scenarios, the crossing probability converged to 1. In all non-crossing scenarios, the crossing probability converged to 0. This meant that the intention recognition method could at least classify pedestrian crossing intention into two different categories. In non-crossing scenarios D, F, and H, pedestrian crossing probabilities were below 0.5 in the whole process, which helped not to mislead the intelligent vehicles. In scenario B, the pedestrian was walking towards the road and suddenly stopped. So, the crossing probability was high at first then decreased quickly, which we thought was reasonable. In crossing scenarios, the intention recognition method could recognize pedestrian crossing intention at least 0.3 s earlier before crossing, which made it possible for intelligent vehicles to decelerate. This conclusion could be explained more clearly in Figure 13. When the crossing intention was convinced, the relative distances between pedestrians and vehicles were still enough for vehicles to decelerate and avoid the collision.

In summary, for the eight defined pedestrian crossing scenarios, the proposed method could recognize pedestrian crossing intention 0.3–0.5 s earlier before crossing and 15 m~20 m away from the theoretical collision position, which provided enough

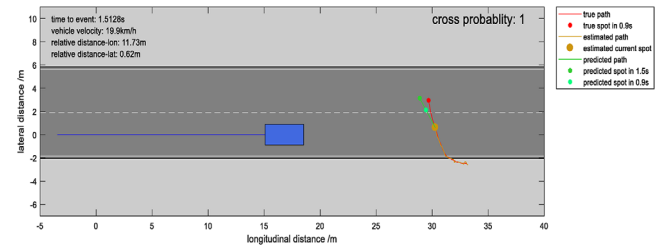


FIGURE 14 Pedestrian trajectory prediction interface

time and distance for intelligent vehicles to conduct the deceleration and obstacle avoidance operation.

4.6 | Validation of pedestrian trajectory prediction

To exhibit the performance of the modified particle filter, a pedestrian trajectory prediction interface was developed as shown in Figure 14. The light grey areas were sidewalks and the dark grey areas were the motorway. The blue rectangle represented the test vehicle and the blue line represented the trajectory of the vehicle. The true path and spot of the pedestrian were highlighted in red. The estimated and predicted objects were shown in brown and green, respectively. The depth of green colour showed the prediction time from now. The lighter the colour, the longer the prediction time. To access the prediction accuracy, the pedestrian trajectory in the future 3 s should be given. However, limited to the data collected, the pedestrian real trajectory in the future 3 s might be lost by the vehicle-mounted sensors. Then the prediction horizon would be less than 3 s.

The trajectory prediction process was continuous. However, limited to the form of the thesis, only typical images in the prediction could be exhibited in the following context. The numbers in the images represented the relative time defined in Figure 12, and the blue points showed the future position of the particles. The prediction horizon was 3 s and the results in the eight typical scenarios would be shown in groups.

As presented in Figures 15 and 16, for simple pedestrian motion, the particle filter got good prediction results in the whole process. In scenario A, the probabilities of crossing in the whole process were above 0.5, so the prediction results were straight lines towards the opposite. In scenarios D, F, and H, the probability of crossing in the whole process was below 0.5, so the predictions were straight lines or points near the edge of the road. The particles surrounded the real trajectory tightly in all these scenarios. And the average prediction suited the real trajectory well.

In scenarios B and C, things were quite different, as shown in Figure 17. Due to the intention recognition model, the particles first overstepped the edge of the road in scenario B. Then, with the convincing of the non-crossing intention, the trajectory prediction results converged to the edge of the road, which was more reasonable in this scenario. In scenario C, things were

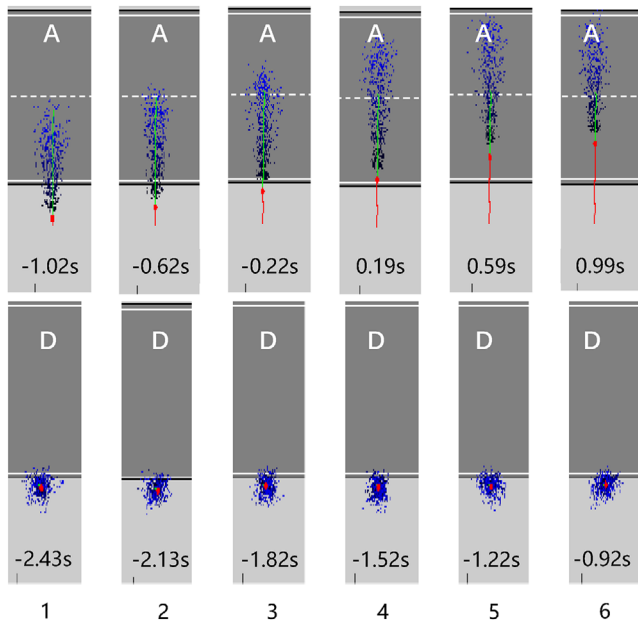


FIGURE 15 Pedestrian trajectory prediction in scenario A, D

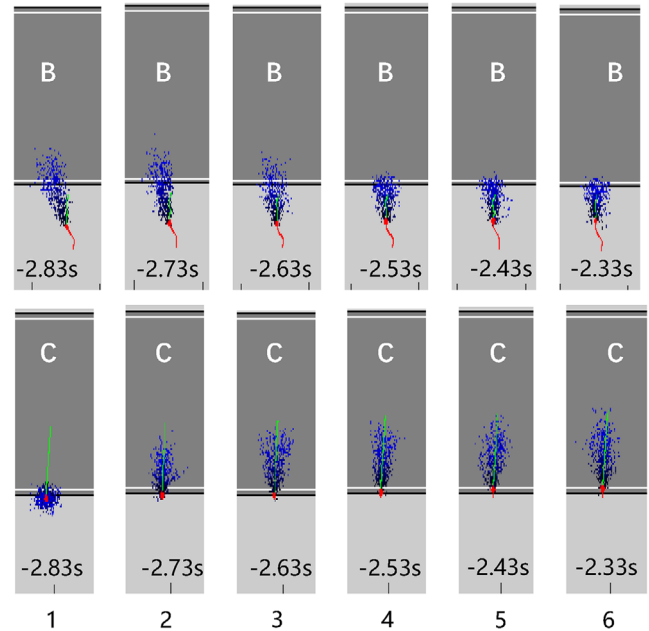


FIGURE 17 17 Pedestrian trajectory prediction in scenario B, C

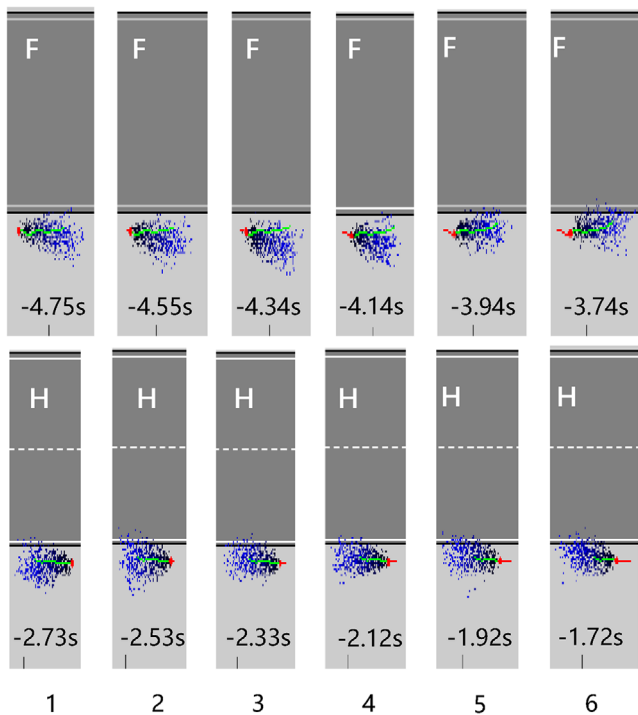


FIGURE 16 Pedestrian trajectory prediction in scenario F, H

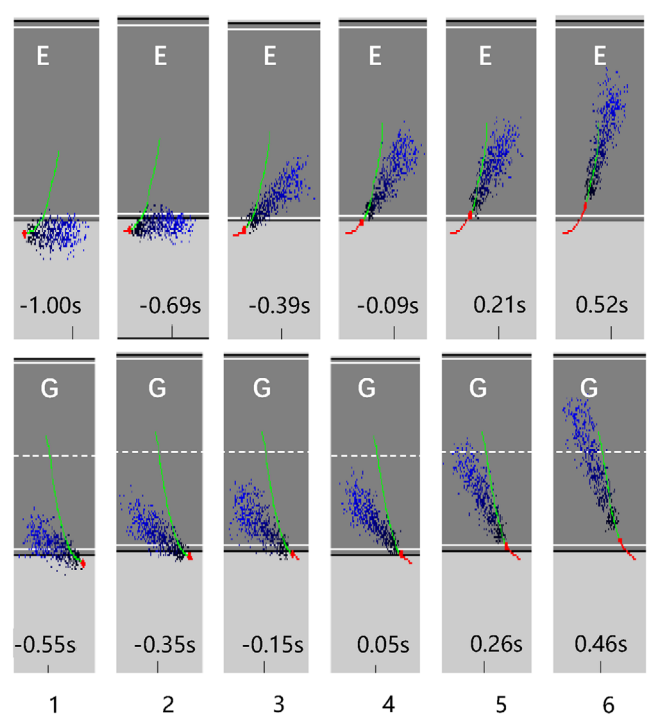


FIGURE 18 Pedestrian trajectory prediction in scenario E, G

similar. The prediction results first converged to the edge of the road then exceeded the edge.

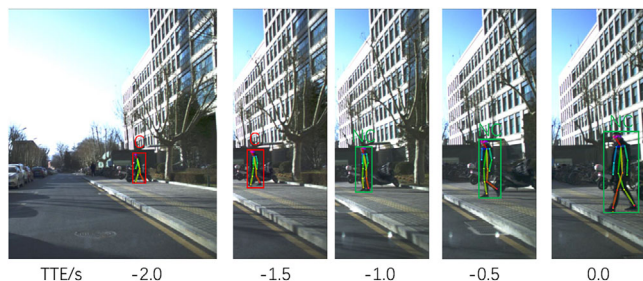
In scenarios E and G, pedestrians crossed the road with turning, as shown in Figure 18. Before the crossing intentions were recognized, the trajectory prediction results were nearly straight lines near the edge of the road. When the intentions were convinced, the trajectories were affected by the orientation of the pedestrian. Due to the neglect of orientation change in the

motion model, the predictions held a relatively large error in these two scenarios. However, in real scenarios, due to pedestrian uncertain motion, the prediction results after pedestrians reached the motorway did not affect the vehicle's decision apparently.

To demonstrate the effectiveness of the proposed trajectory prediction method, we compared our method with the O-LSTM method [44] in the BPI dataset. We reported the

TABLE 9 Trajectory prediction comparison in BPI dataset

Metric	O-LSTM/m	Our Method/m
ADE	0.68	0.25
FDE	0.95	0.43

**FIGURE 19** Results of the intention recognition. Bounding boxes are coloured depending on the presence (red) or absence (green) of C/NC intention as detected by our model

prediction error with two different metrics. Similar to Pellegrini et al. [45] we used:

1. Average displacement error (ADE) - The mean square error (MSE) over all estimated points of a trajectory and the true points. This was introduced in Pellegrini et al. [45].
2. Final displacement error (FDE) - The distance between the predicted final destination and the true final destination at end of the prediction period T_{pred} .

As mentioned, the length of the collected data was mainly 3 – –8 s. And key points of the pedestrian could not be captured when there was a long distance between the pedestrian and vehicle. So, in this experiment, we observed the trajectories for 1.5 s and predicted their paths for the next 1.5 s. We trained the O-LSTM model on 80% samples and tested it on the remaining samples for 20 epochs. The best results are shown in Table 9.

On the one hand, the O-LSTM method could not capture the changes of pedestrian behaviour and its body orientation, which led to greater errors in the short-term trajectory prediction. On the other hand, the O-LSTM method could not recognize pedestrian intention. When the pedestrian walked towards the road curb, the O-LSTM method could not judge if the pedestrian would stop at the road curb. That also led to an increase in the long-term prediction. In summary, compared with the O-LSTM method, the proposed trajectory prediction method based on particle filter can give relatively accurate results in all defined typical scenarios, especially when pedestrian crossing intention or motion changes in the process.

The results of our method could also be illustrated with real images from the camera. We use scenario D as an example. As shown in Figure 19, since the pedestrian is really close to the road curb, it is identified to cross at initial. The predicted trajectory will exceed the road curb at this time. As time goes by, the pedestrian is quickly identified to stop at the road curb. Then,

the predicted trajectory will also be limited within the sidewalk, as shown in Figure 16.

5 | CONCLUSION

In this research, a pedestrian motion prediction method combining Bayesian methods with modified particle filter to avoid pedestrian collisions caused by pedestrian latent crossing behaviour is proposed. This method integrates all the tasks of crossing intention recognition, and trajectory prediction. The main achievements of this paper are as follows:

For pedestrian behaviour (standing, walking, and running) recognition, a behaviour recognition method based on LSTM is proposed as a prior model. Then, the maximum likelihood model and the Bayesian posterior model are proposed to improve the recognition accuracy and eliminate the impact of occlusion. The behaviour recognition model is verified both in the provided BPI dataset and the PIE dataset.

For pedestrian crossing intention recognition, a method based on DBN is proposed. The proposed method can recognize pedestrian crossing intention 0.3–0.5 s earlier before the crossing behaviour. At that time, the distance between pedestrian and vehicle is enough for collision avoidance (15–20 m). For non-crossing scenarios, the method convinces pedestrian intention at least 20 m before the collision.

For pedestrian trajectory prediction, the trajectory prediction method based on a modified particle filter is proposed and the weight matrix is redefined based on the context. The algorithm can use particles to simulate the pedestrian future states combining pedestrian behaviour and intention. The results are compared with the O-LSTM method in the BPI dataset and outperform the baseline in 1.5 s trajectory prediction. Due to the definition of weight, the algorithm can tune its performance dynamically based on the distribution of obstacles and the road curb. The trajectory prediction results provide an important reference for the decision-making and path planning of intelligent vehicles.

For future work, there are still some improvements to be done. For intention recognition, by collecting more data and analysing the internal regulation, the parameters can be optimized using expectation maximization algorithms. For trajectory prediction, it's possible for us to simulate the dynamic motion for the pedestrian and vehicle, then Monte Carlo methods can be used to predict pedestrian future states more accurately.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (52072212, 52072214), National Science Fund for Distinguished Young Scholars (51625503) and Tsinghua-Alliance Joint Research Center for Intelligent Mobility (20193910045).

ORCID

Haoran Wu  <https://orcid.org/0000-0003-2873-4615>

REFERENCES

1. NTSB Office of Safety Recommendations and Communications.: Preliminary Report Released for Crash Involving Pedestrian, Uber Technologies, Inc., Test Vehicle, NTSB News Release, <https://www.ntsbgov/news/press-releases/Pages/NR20180524.aspx> (2018)
2. Schlosser, J., Chow, C.K., Kira, Z.: Fusing LIDAR and images for pedestrian detection using convolutional neural networks. In: Proceedings of 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, pp. 2198–2205 (2016)
3. Haenseler, F.S., et al.: A macroscopic loading model for time-varying pedestrian flows in public walking areas. *Transp. Res. Part. B-Methodol.* 69, 60–80 (2014)
4. Hoogendoorn, S.P., et al.: Continuum theory for pedestrian traffic flow: Local route choice modelling and its implications. *Transp. Res. Part. C-Emerg. Technol.* 59, 183–197 (2015)
5. Yuan, Y., et al.: Macroscopic pedestrian flow simulation using Smoothed Particle Hydrodynamics (SPH). *Transp. Res. Part. C-Emerg. Technol.* 111, 334–351 (2020)
6. Gipps, P.G., Marksjo, B.: A Micro-Simulation Model For Pedestrian Flows. *Math. Comput. Simul.* 27, 95–105 (1985)
7. Okazaki, S.: A study of pedestrian movement in architecture space Part 1: Pedestrian movement by the application on magnetic models. *J. Arch., Plann. Environ. Eng.* 429, 51–59 (1979)
8. Matsushita, S., Okazaki, S.: A study of simulation model for way finding behavior by experiments in mazes. *J. Arch. Plan. Environ. Eng.* 429, 51 (1991)
9. Helbing, D.: A mathematical model for the behavior of pedestrians. *Behav. Sci.* 36(4), 298–310 (1991)
10. Helbing, D.: A fluid-dynamic model for the movement of pedestrians. *Complex Syst.* 6(5), 391–415 (1992)
11. Helbing, D., Vicsek, T.: Optimal self-organization. *New J. Phys.* 1, 13 (1999)
12. Zeng, W., et al.: Application of social force model to pedestrian behavior analysis at signalized crosswalk. *Transp. Res. Part. C-Emerg. Technol.* 40, 143–159 (2014)
13. Zeng, W., et al.: Specification and calibration of a microscopic model for pedestrian dynamic simulation at signalized intersections: A hybrid approach. *Transp. Res. Part. C-Emerg. Technol.* 80, 37–70 (2017)
14. Zhu, Y., Zhao, D., He, H.: Optimal feedback control of pedestrian flow in heterogeneous corridors. *IEEE Trans. Autom. Sci. Eng.* 1–12 (2020). <https://doi.org/10.1109/TASE.2020.2996018>
15. Ning-bo, C., et al.: Destination and route choice models for bidirectional pedestrian flow based on the social force model. *IET Intell. Transp. Syst.* 11(9), 537–545 (2017)
16. Wakim, C.F., Capperon, S., Oksman, J.: A Markovian model of pedestrian behavior. In: Proceedings of 2004 IEEE International Conference on Systems, Man and Cybernetics, The Hague, Netherlands, pp. 4028–4033 (2004)
17. Wakim, C.F., Capperon, S., Oksman, J.: Design of pedestrian detection systems for the prediction of car-to-pedestrian accidents. In: Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems, Washington, WA, pp. 696–701 (2004)
18. Boukas, E., et al.: Robot guided crowd evacuation. *IEEE Trans. Autom. Sci. Eng.* 12(2), 739–751 (2014)
19. Kwak, J.-Y., et al.: Pedestrian's intention prediction based on fuzzy finite automata and spatial-temporal features. *Electron. Imag.* 2016(3), 1–6 (2016)
20. Kwak, J.-Y., Ko, B.C., Nam, J.-Y.: Pedestrian intention prediction based on dynamic fuzzy automata for vehicle driving at nighttime. *Infrared Phys. Technol.* 81, 41–51 (2017)
21. Zhao, J., et al.: Trajectory tracking and prediction of pedestrian's crossing intention using roadside LiDAR. *IET Intell. Transp. Syst.* 13(5), 789–795 (2019)
22. Bonnin, S., et al.: General behavior prediction by a combination of scenario-specific models. *IEEE Trans. Intell. Transp. Syst.* 15(4), 1478–1488 (2014)
23. Peng, S., et al.: Pedestrian motion recognition via Conv-VLAD integrated spatial-temporal-relational network. *IET Intell. Transp. Syst.* 14(5), 92–400 (2019)
24. Hashimoto, Y., et al.: A probabilistic model of pedestrian crossing behavior at signalized intersections for connected vehicles. *Transp. Res. Part. C-Emerg. Technol.* 71, 164–181 (2016)
25. Ravichandar, H.C., Dani, A.P.: Human intention inference using expectation-maximization algorithm with online model learning. *IEEE Trans. Autom. Sci. Eng.* 14(2), 855–868 (2016)
26. Schneider, N., Gavrilu, D.M.: Pedestrian path prediction with recursive bayesian filters: A comparative study. In: Proceedings of the 35th German Conference on Pattern Recognition, Saarbrücken, Germany, pp. 174–183 (2013)
27. Xue, P., et al.: Crossing-road pedestrian trajectory prediction via encoder-decoder LSTM. In: Proceedings of International Conference on Intelligent Transportation Systems, Auckland, New Zealand, pp. 2027–2033 (2019)
28. Goldhammer, M., et al.: Early prediction of a pedestrian's trajectory at intersections. In: Proceedings of International Conference on Intelligent Transportation Systems, The Hague, Netherlands, pp. 237–242 (2013)
29. Iryo-Asano, M., Alhajyaseen, W.: Consideration of a pedestrian speed change model in the pedestrian-vehicle safety assessment of signalized crosswalks. *Transp. Res. Procedia* 21, 87–97 (2017)
30. Keller, C.G., Hermes, C., Gavrilu, D.M.: Will the pedestrian cross? Probabilistic path prediction based on learned motion features. In: Proceedings of Joint Pattern Recognition Symposium, Springer, Berlin, Heidelberg, pp. 386–395 (2011)
31. Keller, C.G., Gavrilu, D.M.: Will the pedestrian cross? A study on pedestrian path prediction. *IEEE Trans. Intell. Transp. Syst.* 15(2), 494–506 (2013)
32. Quintero, R., et al.: Pedestrian path prediction using body language traits. In: Proceedings of IEEE Intelligent Vehicles Symposium, Dearborn, MI, pp. 317–323 (2014)
33. Kooij, J.F.P., et al.: Context-based pedestrian path prediction. In: Proceedings of European Conference on Computer Vision, Zurich, Switzerland, pp. 618–633 (2014)
34. Rehder, E., Kloeden, H.: Goal-directed pedestrian prediction. In: Proceedings IEEE International Conference on Computer Vision Workshop (ICCVW), Santiago, Chile, pp. 50–58 (2015)
35. Karasev, V., et al.: Intent-aware long-term prediction of pedestrian motion. In: Proceedings of IEEE International Conference on Robotics and Automation, Stockholm, Sweden, pp. 2543–2549 (2016)
36. Bandyopadhyay, T., et al.: Intention-aware pedestrian avoidance. *Experimental Robotics*, Springer, Berlin, pp. 963–977 (2013).
37. Ferguson, S., et al.: Real-time predictive modeling and robust avoidance of pedestrians with uncertain, changing intentions. *Algorithmic Foundations of Robotics XI*, Springer, Berlin, pp. 161–177 (2015)
38. OpenPose: Real-time multi-person keypoint detection library for body, face, hands, and foot estimation. <https://github.com/CMU-Perceptual-Computing-Lab/openpose> (2017). Accessed 24 April 2017
39. Hashimoto, Y., et al.: A probabilistic model for the estimation of pedestrian crossing behavior at signalized intersections. In: Proceedings of IEEE 18th International Conference on Intelligent Transportation Systems, Gran Canaria, Spain, pp. 1520–1526 (2015)
40. Yu, D., et al.: Continuous pedestrian orientation estimation using human keypoints. In: Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS), Sapporo, Japan, pp. 1–5 (2019)
41. Minguez, R.Q., et al.: Pedestrian path, pose, and intention prediction through gaussian process dynamical models and pedestrian activity recognition. *IEEE Trans. Intell. Transp. Syst.* 20(5), 1803–1814 (2018)
42. Rasouli, A., Kotseruba, I., Tsotsos, J.K.: Pedestrian action anticipation using contextual feature fusion in stacked rnns. *arXiv preprint arXiv:2005.06582* (2020)
43. Rasouli, A., et al.: Pie: A large-scale dataset and models for pedestrian intention estimation and trajectory prediction. In: Proceedings of

- International Conference on Computer Vision, Seoul, Korea, pp. 6262–6271 (2019)
44. Alahi, A., et al.: Social lstm: Human trajectory prediction in crowded spaces. In: Proceedings of International Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, pp. 961–971 (2016)
 45. Pellegrini, S., et al.: You'll never walk alone: Modeling social behavior for multi-target tracking. In: Proceedings of International Conference on Computer Vision, Kyoto, Japan, pp. 261–268 (2009)
 46. Gordon, N., Salmond, D., Smith, A.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In: IEE Proceedings F (Radar and Signal Processing) 1 40(2), IET Digital Library, pp. 107–113 (1993)

How to cite this article: Xu, Q., et al.: Roadside pedestrian motion prediction using Bayesian methods and particle filter. *IET Intell. Transp. Syst.* 15, 1167–1182 (2021). <https://doi.org/10.1049/itr2.12090>