

# Vulnerable Road User Trajectory Prediction for Autonomous Driving Using a Data-Driven Integrated Approach

Hao Chen<sup>ID</sup>, Yinhua Liu<sup>ID</sup>, Chuan Hu<sup>ID</sup>, and Xi Zhang<sup>ID</sup>, *Senior Member, IEEE*

**Abstract**—In this paper, Vulnerable Road User (VRU) trajectory prediction for autonomous driving based on the Intention-Attention-Gate Recurrent Unit (IA-GRU), Improved Social Force Model (ISFM) and Adaptive Boosting (AdaBoost) is systematically investigated. Firstly, a novel IA-GRU is proposed for VRU (pedestrian, cyclist, and electric cyclist) trajectory prediction. VRU intention (waiting/crossing), VRU heterogeneity (age and gender), VRU-VRU interactions and VRU-dynamic vehicle interactions are taken into account. Attention is used to obtain the influence weights of the above factors used for VRU trajectory prediction. Secondly, a micro-dynamic ISFM is developed for VRU trajectory prediction. The impact of zebra crossing, collision avoidance with vehicles and VRUs, and VRU heterogeneity are considered. Moreover, traffic data collected by an unmanned aerial vehicle (UAV) is obtained and analyzed, and the parameters of the ISFM are calibrated by the Maximum Likelihood Estimation (MLE). Finally, a data-driven integrated approach based on the IA-GRU and ISFM is proposed, and AdaBoost is used to prevent the model from overfitting and improve the prediction accuracy. The results indicate that the integrated model outperforms the existing methods, and the prediction accuracy is improved by more than 11% based on the collected traffic data, which can give us great confidence to use the integrated model in the autonomous driving domain to improve the safety of VRUs.

**Index Terms**—Autonomous driving, data-driven integrated model, VRU trajectory prediction, VRU heterogeneity, VRU intention.

## I. INTRODUCTION

RECENTLY, to reduce the injuries to Vulnerable Road Users (VRUs) caused by traffic accidents, Advanced Driver Assistance System (ADAS) such as Pedestrian Detection System (PDS), Lane Keeping System (LKS), Automatic Emergency Braking System (AEBS) and so on is gaining more and more attentions [1], [2], [3]. At the same time, with

Manuscript received 30 September 2022; revised 16 January 2023; accepted 6 March 2023. Date of publication 20 March 2023; date of current version 7 July 2023. This work was supported in part by the Key Special Projects for International Cooperation in Science and Technology Innovation under Grant 2019YFE0100200 and in part by the National Natural Science Foundation of China under Grant 51875362. The Associate Editor for this article was Y. Huang. (*Corresponding authors: Yinhua Liu; Xi Zhang*)

Hao Chen and Yinhua Liu are with the School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200240, China (e-mail: braver1989@usst.edu.cn; liuyinhua@usst.edu.cn).

Chuan Hu and Xi Zhang are with the School of Mechanical Engineering and the MoE Key Laboratory of Artificial Intelligence, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: chuan.hu@sjtu.edu.cn; braver1980@sjtu.edu.cn).

Digital Object Identifier 10.1109/TITS.2023.3254809

the development of Artificial Intelligence (AI), autonomous driving has received unprecedented attention from scientific research institutes, enterprises and governments [4], [5], [6]. The dynamic changes in the trajectories of the VRUs (pedestrians, cyclists, and electric cyclists) bring great challenges to the decision-making and planning of the autonomous vehicles. Why are autonomous vehicles not widely used at present? One of the key reasons is that the traffic environment is increasing complexity, especially for the VRUs who are with a high degree of uncertainty [7]. Liu et al. [8] provided a survey on dynamic vehicle trajectory prediction, which contains methods using deep-learning approaches. A version of TNT was presented publicly available. A fair comparison of results on public dataset were illustrated, and the accuracy of dynamic vehicle trajectory prediction was improved based on the assistance of the author's work. Compared with the dynamic vehicles, VRUs have higher randomness and uncertainty, they can change the directions and speeds of their movements at any time. Existing studies have shown that a successful 30cm predicted trajectory of VRUs can avoid collisions with vehicles, and emergency braking performs 0.16s in advance can reduce the VRU's injury rate by 50% [9]. Thus, predicting the VRU intention and their trajectories are of great significance for reducing the occurrence of collisions between the autonomous vehicles and VRUs, and they also can improve the driving safety of the autonomous vehicles. Huang et al. [10] presented a survey on road user (pedestrian and vehicle) trajectory-prediction methods. The methods of physics, classic machine learning, deep learning and reinforcement learning were discussed and analyzed in detail, and the performance of each kind of method was evaluated. Leon and Gavrilescu [11] provided a review on tracking and trajectory prediction methods of road users (pedestrian and vehicle), and approaches of deep neural networks and stochastic techniques were discussed. Rudenko et al. [12] proposed a survey on human motion trajectory prediction, and presented a classification based on the motion modeling approaches and the using level of contextual information. The existing datasets and evaluation indicators were introduced and analyzed, the limitations and further research were also discussed. At present, there are two main types of VRU trajectory prediction methods: kinematic models and Neural Networks (NNs). The existing VRUs' kinematic models commonly include Constant Velocity Model (CV), Constant Acceleration

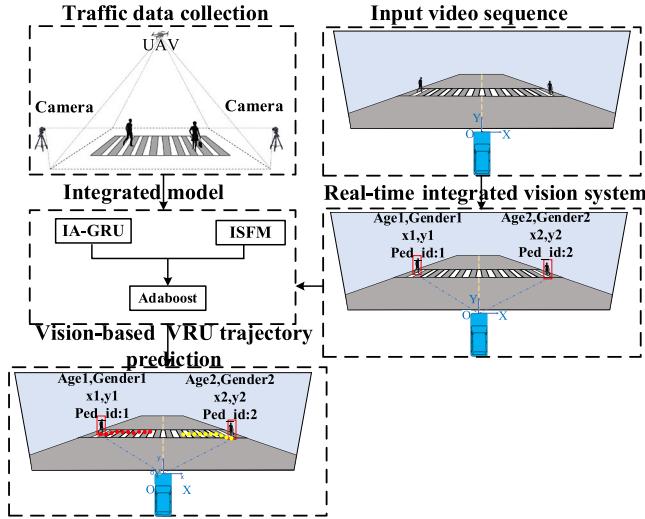


Fig. 1. Procedure of VRU trajectory prediction process.

Model (CA) and Social Force Model (SFM) [13], [14], [15], and VRUs' Neural Network prediction methods commonly include Long Short Term Memory Network (LSTM) and Gate Recurrent Unit (GRU) [16], [17], [18]. The above CV and CA based on the constant parameters are difficult to describe the dynamic changes of pedestrians, and the prediction accuracy needs to be further improved. The above SFM can describe the interactions between road users, and the accuracy of trajectory prediction is higher than those of the CV and CA. However, the disadvantages of SFM are that it needs some rules and features, it can only predict the immediate reaction of the VRUs, and it cannot consider long-term dependence on information like the GRU and LSTM do. Despite GRU and LSTM have shown the advantages in trajectory prediction with relatively high accuracy, they still need some improvements. For example, GRU and LSTM in the trajectory prediction literature cannot describe the VRU's social behavior like the SFM does. Thus, to improve the accuracy of VRU trajectory prediction and to reduce the disadvantages of the SFM and GRU or LSTM, an integrated approach based on the two above methods is required.

In this paper, to ensure the driving safety of the autonomous vehicles, a data-driven integrated approach based on the IA-GRU, ISFM and AdaBoost is developed for the VRU trajectory prediction. The procedure of the VRU trajectory prediction process is shown in Fig. 1, and the main contributions of this paper are summarized as follows:

1) A novel IA-GRU based on the Neural Networks is presented for the VRU trajectory prediction. In the IA-GRU, VRU intention, VRU heterogeneity, VRU-VRU interactions and VRU-dynamic vehicle interactions are considered. Attention is developed to achieve the influence weights of the above influencing factors.

2) A micro-dynamic ISFM is developed for the VRU trajectory prediction. In the ISFM, the influence of zebra crossing, VRU heterogeneity, and collision avoidance with VRUs and vehicles are considered. Compared with our previous work in Zhang et al. [3], more kinds of VRUs are considered,

the generalization ability of the ISFM is stronger, and traffic data is further expanded, which expands the ISFM's practical application in the autonomous driving domain.

3) To combine the complementary advantages of the IA-GRU and ISFM, a data-driven integrated approach is proposed for VRU trajectory prediction. AdaBoost is used to prevent the model from overfitting and improve the prediction accuracy. The model validation results show that the accuracy of the integrated model is higher than the existing models.

The rest of the paper is presented as follows. Section II presents a brief review of the related work for the VRU trajectory prediction. Section III presents a data-driven integrated model based on the IA-GRU and ISFM for the VRU trajectory prediction, traffic data is collected and analyzed, and parameter calibration of the ISFM is conducted. Section IV conducts model validation with the datasets. Section V gives the conclusions of this paper.

## II. RELATED WORK

On one hand, there are many researches on the VRU trajectory prediction using the Neural Networks. Ma et al. [19] proposed a LSTM-based to predict the trajectories of traffic-agents. An instance layer was developed to obtain instances' movements and interactions. The performance of the LSTM-based was compared with the prior prediction methods. The results showed that the proposed model could predict the traffic-agents' trajectories with higher accuracy. Song et al. [20] presented a deep convolutional LSTM to predict pedestrians' trajectories, the tensor and convolution could acquire better pedestrian-environment interactions. The experimental results demonstrated that the proposed model could predict more accurate trajectories for a dense crowd in evacuation and counter-flow. Bhattacharyya et al. [21] developed a Bayesian-LSTM to predict pedestrians' trajectories in 1 second, and particular attention was paid to modeling the uncertainty of natural traffic scenes. The experimental results showed that the proposed model was possible to predict pedestrians' trajectories with significant accuracy. Saleh et al. [22] proposed a B-LSTM to predict VRU's trajectory. The proposed model achieved 2.93 and 4.12 pixels in 2 and 3 seconds, respectively. Moreover, the proposed model was compared with other baseline models, and the lowest margin of average displacement error was achieved. Haddad and Lam [23] proposed a graph-based GRU to achieve interactions between the pedestrians, and the pedestrians' future trajectories and velocities could be predicted. The results showed that the proposed model outperformed the state-of-the-art methods in two public datasets. Zhang and Zheng [24] presented a data-driven pedestrian trajectory predictor called MLP-social-GRU. GRU was used to achieve hidden features of pedestrian motion patterns. The predictor was evaluated on two public datasets, and it was superior than the existing models. The above works show that VRU-VRU interactions and VRU-vehicle interactions affect the accuracy of the VRU trajectory prediction. However, to further expand the application of the VRU trajectory prediction in the autonomous driving domain, and to improve the accuracy of trajectory prediction, not

only various interactions should be considered, but also VRU intention (waiting or crossing), VRU heterogeneity (age and gender) should be considered in the NNs. Moreover, NNs cannot describe VRU's social behavior like the SFM does, and the accuracy of trajectory prediction will also be decreased.

On the other hand, some researchers have conducted the VRU trajectory prediction based on the Social Force Model (SFM). Zhu et al. [25] proposed a social force-based motion model to predict the trajectory of the pedestrians and cyclists, and the conflicts among pedestrians, cyclists, and vehicles were taken into consideration. The experimental results demonstrated that the proposed model could effectively track the pedestrians' movements and predict their trajectories. Liu et al. [26] developed a multi-pedestrian online tracking method based on a social force-predicted deformable keypoints mapping. To deal with the problems of partial occlusion and similar appearance, a SFM was established to predict the trajectories of pedestrians. Yang and Ozguner [27] investigated how pedestrian motion was affected by the surrounding road users, and a SFM was developed. The effectiveness of the proposed model was verified. Chen et al. [28] proposed a MSFM for the pedestrian trajectory prediction, road users, traffic lights, and crosswalk boundary were taken into account. The results showed that the proposed model outperformed the existing models. Zeng et al. [29] proposed a social force-based approach to predict pedestrians' movements at signalized crosswalks, traffic lights, vehicles and pedestrians were considered. Model validation was conducted by comparing the predicted trajectories with the observed trajectories. Huang et al. [30] developed a SFM to analyze the cyclists' behaviors with the heterogeneous traffic. The parameters of the SFM were calibrated by the Maximum Likelihood Estimation (MLE). The results showed that the model could predict cyclists' crossing behaviors as in the real world. The above works show that the SFM can take the VRU-VRU interactions, VRU-vehicle interactions and transportation facilities into consideration for the VRU trajectory prediction. However, most researchers just consider that each VRU has the same degree of influence in the SFM, VRU heterogeneity (age and gender) which is a key factor for improving the prediction accuracy is not considered. In addition, the SFM only considers the instantaneous information transfer, and fails to incorporate the pedestrians' behaviors over a period of time like the LSTM/GRU does, therefore, resulting in reducing the accuracy of predicted trajectories. To make the NNs and SFM complement each other's advantages, and to improve the accuracy of the VRU trajectory prediction, an integrated approach is needed.

Different from the previous works above, in this paper, a data-driven integrated approach based on the IA-GRU and ISFM is proposed for the VRU trajectory prediction, and AdaBoost is used to prevent the model from overfitting and improve the prediction accuracy. Compared with the LSTM, the GRU has similar prediction accuracy, but the GRU has fewer parameters, simpler network structures, and faster convergence speed, so the GRU is selected for the VRU trajectory prediction. VRU heterogeneity (age and gender), VRU-VRU interactions and VRU-dynamic vehicle interactions

TABLE I  
MULTI-FEATURE VARIABLES AND VARIABLE CODING

| Variable coding | VRU heterogeneity |                    | Vehicle      |
|-----------------|-------------------|--------------------|--------------|
|                 | Gender            | Age type           | Vehicle type |
| 1               | Male              | Young(18-30)       | Car          |
| 2               | Female            | Middle-aged(30-55) | Medium car   |
| 3               | /                 | Old(>55)           | Large car    |

are considered in both IA-GRU and ISFM. VRU intention and Attention are also considered in the IA-GRU. In addition, the influence of zebra crossing is considered in the ISFM. Traffic data is collected by an unmanned aerial vehicle (UAV), and the non-measurable parameters of the ISFM are conducted based on the MLE method. The performance of the data-driven integrated model is compared with the existing VRU trajectory prediction methods. The results show that the integrated model can outperform the existing models with significant accuracy, which can give us great feasibility for improving the safety of VRUs in the autonomous driving domain.

### III. INTEGRATED MODEL BASED ON IA-GRU AND ISFM FOR VRU TRAJECTORY PREDICTION

#### A. IA-GRU for VRU Trajectory Prediction

It is supposed that all the VRUs choose to cross the zebra crossing when the surrounding traffic environment is safe, and choose to wait when the environment is unsafe. However, some VRUs may respond to the unsafe environment with different strategies, leading to the possibility of potential conflicts with the vehicles [31]. For example, certain VRUs may cross the zebra crossing and ignore the risk of conflicting with the approaching vehicles, which may threaten their own safety. In other words, to help the autonomous vehicles better understand the complex traffic environment, it is very important for the autonomous vehicles to recognize the waiting/crossing intention of VRUs when they are crossing the zebra crossing.

In this paper, an un-signalized zebra crossing is focused on, and a multi-feature-based VRU intention model is developed to predict the VRU's waiting/crossing intention. Table I shows the multi-feature variables and variable coding used in the VRU intention model. It includes one binary variable and two tri-class variables. Distance and vehicle speed which are two continuous variables are calculated in real-time based on the traffic environment. Distance is defined as the distance to the potential collision point if the dynamic vehicle keeps the current direction. The framework of the VRU intention prediction is shown in Fig. 2. The multi-feature sequence  $I$  is used as the input of the intention prediction network, which can be shown as follows:

$$I = (I_1, I_2, \dots, I_n) \quad (1)$$

$$I_t = (O_{i1}, O_{i2}, \dots, O_{i(t-1)}, O_{it}) \quad (2)$$

$$O_{it} = (M_{i1t}, M_{i2t}, M_{i3t}, M_{i4t}, M_{i5t}) \quad (3)$$

where  $M_{i1t} \dots M_{i5t}$  are the multi-feature variables (gender type, age type, distance, vehicle speed, vehicle type) at the

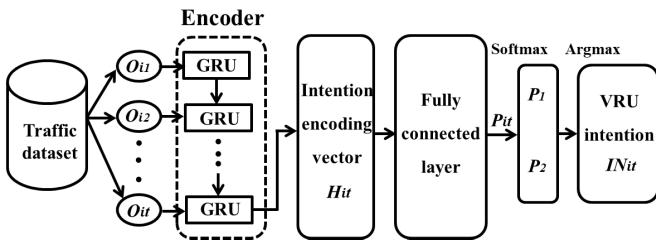


Fig. 2. Framework of the VRU intention prediction.

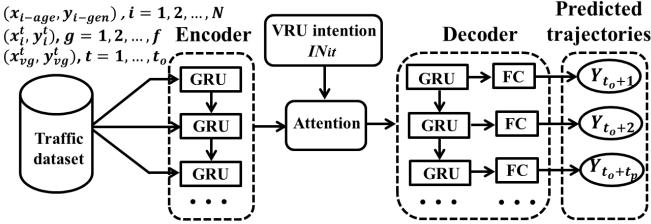


Fig. 3. Framework of the IA-GRU for the VRU trajectory prediction.

time  $t$ ;  $O_{it}$  is a collection of VRU  $i$ 's multi-feature variables at the time  $t$ ;  $n$  is the number of VRUs.

The multi-feature sequence  $I_i$  can obtain the intention encoding vector  $H_{it}$  based on the Encoder. At the time  $t$ , the intention encoding vector  $H_{it}$  is passed through the fully connected layer. The softmax function is used to obtain the intent prediction probability distribution  $P_{it}$ . The  $P_{it}$  is input to the argmax function to achieve the VRU intention  $IN_{it}$ , which can be expressed as follows:

$$H_{it} = GRU(H_{i(t-1)}, O_{it}; W_{s1}) \quad (4)$$

$$P_{it} = softmax(FC(H_{it}; W_{s2}); W_{s3}) \quad (5)$$

$$IN_{it} = argmax(P_{it}; W_{s4}) \quad (6)$$

where  $W_{s1}$  is the parameter of Encoder;  $FC$  is a fully connected layer;  $W_{s2}$  is the parameter of  $FC$ ;  $W_{s3}$  is the parameter of the softmax function;  $W_{s4}$  is the parameter of the argmax function.

In this paper, a novel IA-GRU based on the NNs is proposed for the VRU trajectory prediction, and in the IA-GRU, Attention is used to obtain the influence weights of the VRU intention, VRU heterogeneity, VRU-VRU interactions, and VRU-dynamic vehicle interactions. The framework of the IA-GRU is shown in Fig. 3. Three main parts are included: Encoder, Attention, and Decoder.  $X$  is all the VRUs' trajectories,  $V$  is all the dynamic vehicles' trajectories, and  $L$  is all the VRUs' age and gender type, which can be shown as follows:

$$X = (X_1, X_2, \dots, X_N) \quad (7)$$

$$X_i = \{(x_i^t, y_i^t), t = 1, \dots, t_o + t_p, i = 1, \dots, N\} \quad (8)$$

$$V = (V_1, V_2, \dots, V_f) \quad (9)$$

$$V_g = \{(x_{vg}^t, y_{vg}^t), t = 1, \dots, t_o, g = 1, \dots, f\} \quad (10)$$

$$L = (L_1, L_2, \dots, L_N) \quad (11)$$

$$L_i = \{(x_{i-age}, y_{i-gen}), i = 1, \dots, N\} \quad (12)$$

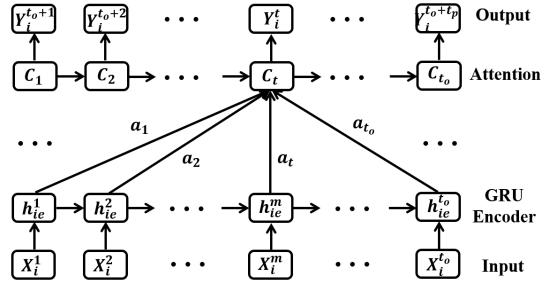


Fig. 4. Framework of the Attention for the VRU trajectory prediction.

where  $N$  is the number of all the VRUs;  $f$  is the number of all the dynamic vehicles;  $t_o$  is the observed time;  $t_p$  is the predicted time;  $(x_i^t, y_i^t)$  is VRU  $i$ 's trajectory points at the time  $t$ ;  $(x_{vg}^t, y_{vg}^t)$  is dynamic vehicle  $g$ 's trajectory points at the time  $t$ ;  $(x_{i-age}, y_{i-gen})$  is VRU  $i$ 's age and gender type.

1) *Encoder*: In this part, sequences of trajectories of the VRU  $i$  ( $X_i^t$ ) and dynamic vehicle  $g$  ( $V_g^t$ ), and VRU heterogeneity ( $L_i$ ) are entered into the Encoder layer. The hidden state ( $h_{ie}^t$ ) is achieved, which can be expressed as follows:

$$\begin{cases} h_{ie}^t = GRU(h_{ie}^{t-1}, X_i^t, V_g^t, L_i; W_{e1}) \\ H_{ie} = (h_{ie}^1, h_{ie}^2, \dots, h_{ie}^t) \end{cases} \quad (13)$$

where  $h_{ie}^t$  is the hidden state of the VRU  $i$  at the time  $t$ ;  $t = 1, \dots, t_o$ ;  $W_{e1}$  is the parameter of the Encoder;  $H_{ie}$  is the collection of the hidden state sequences of the VRU  $i$ .

2) *Attention*: As the VRU  $i$ 's future trajectories are not only influenced by the previous hidden state  $h_{ie}^{t-1}$  and current trajectory  $X_i^t$ , but also influenced by the VRU intention, VRU heterogeneity, VRU-VRU interactions, and VRU-dynamic vehicle interactions. In this part, to obtain VRU  $i$ 's more comprehensive movement trend information, Attention is used to obtain the influence weights of the above influencing factors. Five features are selected to express the influence on the movement of the VRU  $i$ , which can be shown as follows:

$$D_{ij}^t = \left\{ ((x_i^t - x_j^t)^2 + (y_i^t - y_j^t)^2 \right\}^{1/2} \quad (14)$$

$$v_{ij}^t = \left\{ (\dot{x}_i^t - \dot{x}_j^t), (\dot{y}_i^t - \dot{y}_j^t) \right\} \quad (15)$$

$$D_{ivg}^t = \left\{ ((x_i^t - x_{vg}^t)^2 + (y_i^t - y_{vg}^t)^2 \right\}^{1/2} \quad (16)$$

$$\theta_{ij}^t = [D_{ij}^t, v_{ij}^t, D_{ivg}^t, L_i, IN_{it}] \in R^5 \quad (17)$$

$$\theta_i = [\theta_{i1}, \theta_{i2}, \dots, \theta_{in}] \quad (18)$$

where  $D_{ij}^t$  is the distance between the subject VRU  $i$  and surrounding VRU  $j$  at the time  $t$ ;  $v_{ij}^t$  is the relative velocity between the subject VRU  $i$  and surrounding VRU  $j$  at the time  $t$ ;  $D_{ivg}^t$  is the distance between the subject VRU  $i$  and dynamic vehicle  $g$  at the time  $t$ ;  $\theta_{ij}^t$  is the collection of five features at the time  $t$ ;  $\theta_i$  is the collection of  $\theta_{ij}^t$ ;  $t = 1, \dots, t_o$ .

The shorter the distance of the  $D_{ij}^t$  and  $D_{ivg}^t$ , the greater the impact on the movement of the subject VRU  $i$ . Fig. 4 shows the framework of the Attention. The tensor  $C_t$  that the subject VRU  $i$  needs to make a movement decision is summarized,

which can be shown as follows:

$$P_i = FC(\theta_i; W_{e2}) \quad (19)$$

$$H_e^{t_o} = [H_{1e}^{t_o}, H_{2e}^{t_o}, \dots, H_{ne}^{t_o}] \quad (20)$$

$$e_t = FC(P_i, H_e^{t_o}; W_{e3}) \quad (21)$$

$$a_t = \exp(e_t) / \sum_{q=1}^{t_o} \exp(e_q) \quad (22)$$

$$C_t = \sum_{t=1}^{t_o} a_t H_e^{t_o} \quad (23)$$

where  $FC$  is a fully connected layer;  $W_{e2}$  is the parameter of  $FC$ ;  $P_i$  is the feature vector of the VRU  $i$ ;  $W_{e3}$  is the parameter of the  $FC$ ;  $e_t$  is the importance of the  $P_i$  and  $H_e^{t_o}$  to the  $C_t$ ;  $a_t$  is the contribution weight of the features.

3) *Decoder*: In this part, the hidden state of the Decoder is initialized as the  $h_{id}^{t_0}$ . The tensor  $G_i^{t_0}$  shows the influence of the VRU intention, VRU heterogeneity, VRU-VRU interactions, and VRU-dynamic vehicle interactions. Moreover, the hidden state  $h_{id}^{t_0}$  can be achieved with the Gaussian noise  $z$  and tensor  $G_i^{t_0}$ . To transform the two-dimensional trajectory points  $Y_i^{t-1}$  into a feature vector, a fully connect network is used. The hidden state  $h_{id}^t$  can be obtained with the feature vector based on the GRU. The predicted trajectories  $Y_i^t$  of the VRU  $i$  can be obtained with the hidden state  $h_{id}^t$  through a fully connect network. The IA-GRU module of Decoder can be expressed as follows:

$$G_i^{t_0} = FC(C_i, h_{id}^{t_0}; W_{e4}) \quad (24)$$

$$h_{id}^{t_0} = [G_i^{t_0}, z] \quad (25)$$

$$h_{id}^t = GRU(h_{id}^{t-1}, FC(Y_i^{t-1}; W_{e5}); W_{e6}) \quad (26)$$

$$Y_i^t = (x_i^t, y_i^t) = FC(h_{id}^t; W_{e7}) \quad (27)$$

where  $W_{e4}$ ,  $W_{e5}$  and  $W_{e7}$  are the parameters of the  $FC$ ;  $W_{e6}$  is the parameter of the Decoder;  $t = t_o + 1, \dots, t_o + t_p$ .

### B. ISFM for VRU Trajectory Prediction

Inspired by our previous work in Zhang et al. [3], a microdynamic ISFM is developed for the VRU trajectory prediction. In the ISFM, the influence of VRU heterogeneity, zebra crossing, VRU-VRU interactions, and VRU-dynamic vehicle interactions are taken into account. In this paper, refer to our previous work in Zhang et al. [3], the composition of the ISFM is described briefly, and the ISFM contains four parts: self-driving force  $\vec{F}_d$  from the destination, repulsive force  $\vec{F}_{ij}$  from the other conflicting VRUs, repulsive/accelerated force  $\vec{F}_{iv}$  from the conflicting vehicles, and repulsive/attractive force  $\vec{F}_{iz}$  from the zebra crossing boundary. The combined force  $\vec{F}_{ic}$  of all the forces can be shown as follows:

$$\vec{F}_{ic} = \vec{F}_d + \vec{F}_{ij} + \vec{F}_{iv} + \vec{F}_{iz} \quad (28)$$

The unit vector  $\vec{e}_o$  of desired direction can be obtained with the VRU's observed start and end trajectory vector ( $\vec{P}_s$  and  $\vec{P}_e$ ) at a time horizon of 1 second. Considering that the predicted trajectory of the VRU is at a time horizon of 2 seconds, the destination trajectory vector  $\vec{P}_d$  can be shown as follows, which is the end trajectory vector  $\vec{P}_e$  moving with

the individual desired speed  $V_d$  and the unit vector  $\vec{e}_o$  after 3 seconds.

$$\begin{cases} \vec{e}_o = \frac{\vec{P}_e - \vec{P}_s}{\|\vec{P}_e - \vec{P}_s\|} \\ \vec{P}_d = \vec{P}_e + 3V_d \vec{e}_o \end{cases} \quad (29)$$

The self-driving force  $\vec{F}_d$  from the destination of the VRU  $i$  can be shown as follows:

$$\vec{F}_d = \frac{(V_d \vec{e}_d - \vec{V}_i)}{\tau_i} \quad (30)$$

where  $\vec{e}_d$  is the unit vector;  $\vec{V}_i$  is the actual speed of the VRU  $i$ ;  $\tau_i$  is the relaxation time of the VRU  $i$ .

The repulsive force  $\vec{F}_{ij}$  from the other conflicting VRUs can be expressed as follows:

$$\begin{cases} \vec{F}_{ij} = \sum_{i=1}^n A_{ij} \exp\left(\frac{(r_{ij} - b_i)}{B_{ij}}\right) \vec{n}_{ji} \\ b_i = \frac{1}{2} \sqrt{\left(\|\vec{P}_j - \vec{P}_i\| + \|\vec{P}_j + \vec{V}_j \Delta t - \vec{P}_i\|\right)^2 - \left(\|\vec{V}_j \Delta t\|\right)^2} \end{cases} \quad (31)$$

where  $A_{ij}$  and  $B_{ij}$  are the strength coefficients of the other conflicting VRUs;  $\vec{n}_{ji}$  is the normalized direction vector;  $r_{ij}$  is the sum of the VRUs' radius;  $b_i$  is the distance between the subject VRU  $i$  and other conflicting VRU  $j$ ;  $\Delta t$  is the time step (0.2 seconds).

The repulsive/accelerated force  $\vec{F}_{iv}$  from the conflicting vehicles can be shown as follows:

$$\vec{F}_{iv} = \begin{cases} \vec{F}_{iv}^r = A_{iv}^r \exp\left(\frac{(r_{iv} - d_{in})}{B_{iv}^r}\right) \vec{n}_{vi}, & \text{if } \vec{V}_i * \vec{n}_{vi} < 0 \\ \vec{F}_{iv}^a = A_{iv}^a \exp\left(\frac{(r_{iv} - d_{iv})}{B_{iv}^a}\right) \vec{e}_d, & \text{otherwise} \end{cases} \quad (32)$$

where  $A_{iv}^r$ ;  $B_{iv}^r$ ;  $A_{iv}^a$  and  $B_{iv}^a$  are the strength coefficients of the conflicting vehicles;  $r_{iv}$  is the sum of the radius of the conflicting vehicle and VRU  $i$ ;  $d_{in}$  is the distance between the nearest vertex of the conflicting vehicle and subject VRU  $i$ ;  $\vec{n}_{vi}$  is the normalized direction vector;  $d_{iv}$  is the distance between the conflicting vehicle and subject VRU  $i$ .

The repulsive/attractive force  $\vec{F}_{iz}$  from the zebra crossing boundary can be shown as follows:

$$\vec{F}_{iz} = \begin{cases} \vec{F}_{iz}^r = A_{iz}^r \exp\left(\frac{(r_i - d_{iz})}{B_{iz}^r}\right) \vec{n}_{zi}, & \text{if VRU } i \\ & \text{is inside} \\ \vec{F}_{iz}^a = A_{iz}^a \exp\left(\frac{-(r_i - d_{iz})}{B_{iz}^a}\right) \vec{n}_{iz}, & \text{otherwise} \end{cases} \quad (33)$$

where  $A_{iz}^r$ ;  $B_{iz}^r$ ;  $A_{iz}^a$  and  $B_{iz}^a$  are the strength coefficients of the zebra crossing boundary;  $r_i$  is the subject VRU  $i$ 's radius;

$d_{iz}$  is the distance between the zebra crossing boundary and the subject VRU  $i$ ;  $\vec{n}_{zi}$  and  $\vec{n}_{iz}$  are the normalized direction vector.

As the combined force  $\vec{F}_{ic}$  of all the forces have been shown above, the combined speed vector  $\vec{V}_{ic}(t_k)$  can be shown as follows:

$$\vec{V}_{ic}(t_k) = \vec{V}_i(t_k) + \vec{F}_{ic}(t_k) \Delta t \quad (34)$$

where  $\vec{V}_i(t_k)$  is the subject VRU  $i$ 's current speed vector at the time  $t_k$ .

Then the VRU  $i$ 's next prediction trajectory is shown as:

$$\vec{P}_{ic}(t_{k+1}) = \vec{P}_i(t_k) + \vec{V}_i(t_k) \Delta t + \frac{1}{2} \vec{F}_{ic}(t_k) \Delta t^2 \quad (35)$$

where  $\vec{P}_i(t_k)$  represents the subject VRU  $i$ 's current trajectory vector at the time  $t_k$ .

### C. Integrated Model for VRU Trajectory Prediction

The integrated methodology can be seen as an ensemble method. Compared with a single model, the integrated model can combine the complementary advantages of the single model, and can obtain more accurate predicted trajectories. In this paper, a data-driven integrated model based on the IA-GRU and ISFM is proposed. Fig. 5 shows the process of VRU trajectory prediction based on the integrated framework. It can be seen that the process of the integrated model consists of two parts. The first part is to predict the VRUs' trajectories using the IA-GRU and ISFM, which are considered as the base models, respectively. To avoid over-fitting, a process of five-fold-cross-training is proposed for each base model, and new training data and new testing data are generated. The five orange and one yellow colour blocks are the VRUs' predicted trajectories based on the IA-GRU, which are used as the new training data and new testing data, respectively. The five red and one dark red colour blocks are the VRUs' predicted trajectories based on the ISFM, which are used as the training data and testing data, respectively. The second part is to predict the VRUs' trajectories using the AdaBoost based on the above data, and the above data are merged into column vectors as the new training data and testing data, respectively, which is used for training and testing the AdaBoost. Finally, the predicted trajectories can be achieved with relatively lower generalization errors. Here the AdaBoost is used as the meta-level model to integrate the IA-GRU and ISFM. It is considered as an improved linear regression model, which can be expressed as follows.

$$\hat{y}(t) = \hat{y}(t-1) + \sum_{i=1}^n w_i * \text{basemodel}_i(t) \quad (36)$$

where  $w_i$  represents the weight coefficient of the base model  $i$ ;  $n$  represents the number of the base models;  $\text{basemodel}_i(t)$  represents the predicted trajectories of the base model  $i$ ;  $\hat{y}(t)$  represents the final predicted trajectories.

The AdaBoost was first introduced by Freund and Schapire [32], which was considered as a classic algorithm for the Boosting. The AdaBoost adopts the weighted majority voting method. By increasing the weights of the base models

- Predicted trajectories based on IA-GRU used as training data
- Predicted trajectories based on IA-GRU used as testing data
- Predicted trajectories based on ISFM used as training data
- Predicted trajectories based on ISFM used as testing data

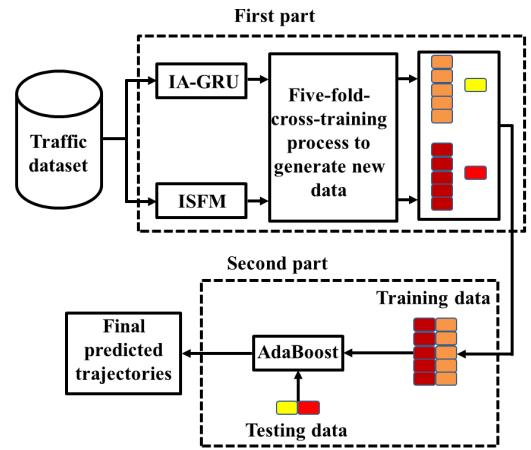


Fig. 5. Process of the VRU trajectory prediction based on the integrated framework.

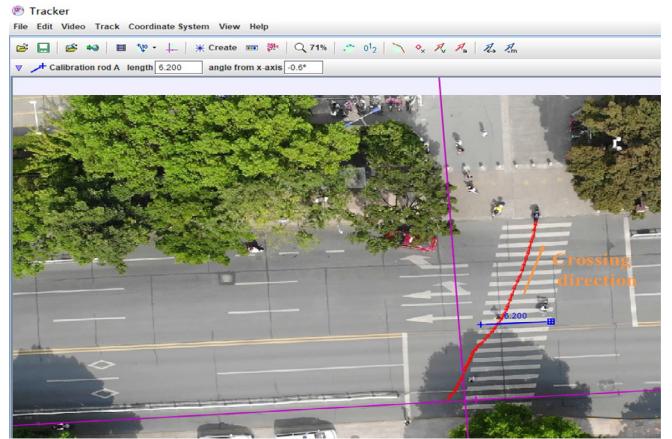


Fig. 6. Diagram of the VRU's trajectory annotation.

with smaller errors, by reducing the weights of the base models with larger errors, and thereby adjusting the weights of each base model in the combined results [33]. The AdaBoost can not only prevent the model from overfitting, but also can improve the robustness of the model. In fact, the AdaBoost is an improved version of the Boost, and it can be easily applied with high accuracy, robustness and stability [34], [35]. To calculate the difference between the predicted trajectories and the actual trajectories of VRUs, the  $L(SF)$  is treated as the loss function of the AdaBoost, and it can be expressed as follows:

$$L(SF) = \min \sum_{i=1}^N L(Y_i, \sum_{m=1}^M \beta_m b(X_i; r_m)) \quad (37)$$

where  $b(X_i; r_m)$  is the basis function;  $\beta_m$  is the coefficient of the basis function;  $r_m$  is the parameter of the basis function;  $N$  is the number of all the VRUs;  $M$  is the number of base models;  $X_i$  is the VRU's actual trajectories;  $Y_i$  is the VRU's predicted trajectories.

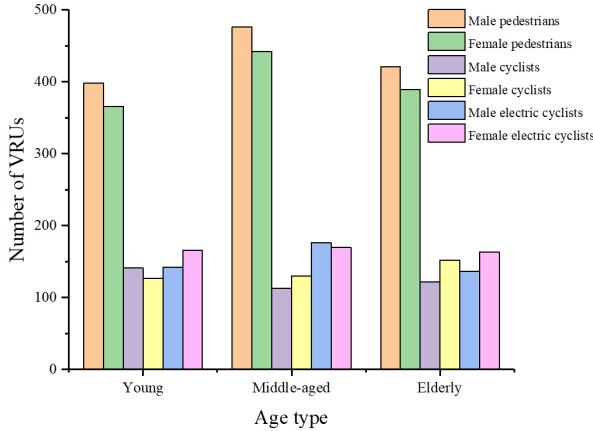


Fig. 7. The number of the VRUs by age and gender type.

TABLE II  
VALUES OF DESIRED SPEED AND RELAXATION  
TIME OF VRUS BY AGE TYPE

| Age type    | Desired speed $V_d$ (m/s) |      |      | Relaxation time $\tau_i$ (s) |      |      |
|-------------|---------------------------|------|------|------------------------------|------|------|
|             | P                         | C    | E    | P                            | C    | E    |
| Young       | 1.45                      | 3.96 | 5.05 | 1.16                         | 1.46 | 1.64 |
| Middle-aged | 1.36                      | 3.52 | 4.66 | 1.45                         | 1.69 | 1.86 |
| Elderly     | 1.04                      | 2.88 | 4.03 | 2.08                         | 1.95 | 2.19 |

#### D. Traffic Data Collection and Parameter Calibration

1) *Traffic Data Collection:* In this paper, an UAV and two cameras were used to collect the traffic data. The region is an un-signaled zebra crossing, and the address is on Bo Le Road, Shanghai. To obtain the VRU heterogeneity, the two cameras were installed on each side of the zebra crossing, and a total of fourteen hours of the traffic data was collected. Fig. 6 shows the VRU's trajectory annotation based on the Tracker software. In total, the trajectories of 2492 pedestrians, 785 cyclists, 953 electric-cyclists and 586 dynamic vehicles were annotated.

Fig. 7 shows the number of the VRUs by age and gender type. Additionally, the amount of traffic data was expanded by rotating [36], and the original trajectories of the VRUs and dynamic vehicles were rotated every 15 degrees along each center. The dataset was enlarged 24 times, and all the generated trajectories were checked to ensure that the kinematic features were reasonable. Finally, a total of 101520 suitable VRUs' trajectories and 14064 suitable vehicles' trajectories were achieved.

2) *Parameter Calibration of ISFM:* In the ISFM, the measurable parameters of the desired speed  $V_d$  and relaxation time  $\tau_i$  based on the VRU heterogeneity are needed, which can be obtained from the collected traffic data. Through the statistics and analysis of the relevant data, it can be found that the values of the desired speed  $V_d$  and relaxation time  $\tau_i$  of the same age type are very close. Thus, to simplify the calculation of the predicted process, the values of the desired speed  $V_d$  and relaxation time  $\tau_i$  of the VRUs by age type are shown in

TABLE II,  $P$  represents pedestrian;  $C$  represents cyclist and  $E$  represents electric cyclist.

A MLE is used to calibrate the non-measurable parameters of the ISFM, and the combined likelihood  $L(\theta_{VRU})$  at each time step is shown as follows:

$$L(\theta_{VRU}) = \prod_{i=1}^n \prod_{k=1}^L \frac{1}{2\pi |\sum|^{1/2}} e^{-\frac{(\vec{F}_{ic}(t_k) - \vec{a}(t_k))^T \sum^{-1} (\vec{F}_{ic}(t_k) - \vec{a}(t_k))}{2}} \quad (38)$$

$$\theta_{VRU} = [A_{ij}, B_{ij}, A_{iv}^r, B_{iv}^r, A_{tv}^a, B_{tv}^a, A_{iz}^r, B_{iz}^r] \quad (39)$$

where  $\vec{a}(t_k)$  is the acceleration of the VRU  $i$  at the time  $t_k$ ;  $\vec{F}_{ic}(t_k)$  is the combined force of the VRU  $i$  at the time  $t_k$ ;  $\theta_{VRU}$  is the collection of the non-measurable parameters, which contains 36 non-measurable parameters based on pedestrian, cyclist and electric cyclist;  $n$  is the number of the VRUs;  $L$  is the number of the time steps.

To simplify the calculation of the calibration process, both sides of equation (38) are taken the logarithm, which can be expressed as follows:

$$\begin{aligned} \ln L(\theta_{VRU}) = & -L_1 L_2 \dots L_n \ln(2\pi) - \frac{L_1 L_2 \dots L_n}{2} \\ & \times \ln(|\sum|) - \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^{L_i} (\vec{F}_{ic}(t_k) - \vec{a}(t_k))^T \\ & \times \sum_{k=1}^{L_i} (\vec{F}_{ic}(t_k) - \vec{a}(t_k)) \end{aligned} \quad (40)$$

where  $L_i$  is the time step of each VRU.

When the  $L(\theta_{VRU})$  is maximized, then the non-measurable parameters of the ISFM can be estimated. In other words, when the negative log-likelihood ( $-\ln L(\theta_{VRU})$ ) is minimized, the non-measurable parameters can be estimated. The minimum of the negative log-likelihood ( $-\ln L(\theta_{VRU})$ ) can be achieved based on the 'fminunc' function in Matlab. In this paper, 70000 trajectories of the VRUs were selected randomly to calibrate the non-measurable parameters  $\theta_{VRU}$ , and the results are presented in TABLE III,  $P$  represents pedestrian;  $C$  represents cyclist;  $E$  represents electric cyclist;  $PC$  represents pedestrian-cyclist;  $PE$  represents pedestrian-electric cyclist;  $CE$  represents cyclist-electric cyclist. It can be seen that the value of strength coefficient  $A_{iv}^r$  is the largest, implying the highest sensitivity of the VRUs to avoid collisions with the vehicles.

#### IV. MODEL VERIFICATION WITH DATASETS

In this paper, 70000 trajectories of the VRUs were selected to train and verify the IA-GRU, and 25000 trajectories of the VRUs were selected to train and test the integrated model. The remaining 6520 trajectories of the VRUs were used to validate the accuracy of the integrated model. In addition, to evaluate the performance of the integrated model, Average Displacement Error (ADE) and Final Displacement Error (FDE) are presented [22], which can be shown as

TABLE III  
CALIBRATION RESULTS OF  $\theta_{VRU}$

| Parameters             | Estimates | P-value   | Equation |
|------------------------|-----------|-----------|----------|
| $A_{ij}/B_{ij}(P)$     | 0.52/2.18 | 0.04/0.03 | (31)     |
| $A_{iv}^r/B_{iv}^r(P)$ | 2.65/2.10 | 0.03/0.02 | (32)     |
| $A_{iv}^a/B_{iv}^a(P)$ | 4.49/1.44 | 0.02/0.03 | (32)     |
| $A_{iz}^r/B_{iz}^r(P)$ | 1.60/3.15 | 0.01/0.03 | (33)     |
| $A_{iz}^a/B_{iz}^a(P)$ | 0.45/2.92 | 0.04/0.03 | (33)     |
| $A_{ij}/B_{ij}(C)$     | 0.45/2.12 | 0.02/0.04 | (31)     |
| $A_{iv}^r/B_{iv}^r(C)$ | 2.54/2.03 | 0.01/0.03 | (32)     |
| $A_{iv}^a/B_{iv}^a(C)$ | 4.36/1.35 | 0.03/0.02 | (32)     |
| $A_{iz}^r/B_{iz}^r(C)$ | 1.39/2.81 | 0.02/0.03 | (33)     |
| $A_{iz}^a/B_{iz}^a(C)$ | 0.51/2.78 | 0.03/0.04 | (33)     |
| $A_{ij}/B_{ij}(E)$     | 0.61/2.01 | 0.03/0.04 | (31)     |
| $A_{iv}^r/B_{iv}^r(E)$ | 2.36/2.27 | 0.03/0.01 | (32)     |
| $A_{iv}^a/B_{iv}^a(E)$ | 4.71/1.33 | 0.02/0.03 | (32)     |
| $A_{iz}^r/B_{iz}^r(E)$ | 1.53/2.82 | 0.03/0.02 | (33)     |
| $A_{iz}^a/B_{iz}^a(E)$ | 0.54/2.85 | 0.04/0.02 | (33)     |
| $A_{ij}/B_{ij}(PC)$    | 1.16/2.21 | 0.04/0.02 | (31)     |
| $A_{ij}/B_{ij}(PE)$    | 1.25/2.34 | 0.02/0.03 | (31)     |
| $A_{ij}/B_{ij}(CE)$    | 1.38/2.16 | 0.01/0.04 | (31)     |

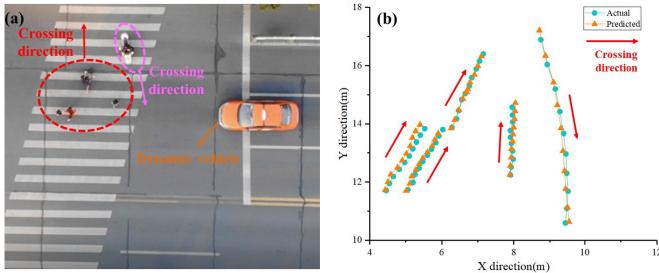


Fig. 8. VRUs' predicted trajectories based on the integrated model.

follows:

$$ADE = \frac{\sum_{i=1}^N \sum_{t=t_0+1}^{t=t_0+t_p} \left\| \vec{P}_{ip}^t - \vec{P}_{ia}^t \right\|_2}{N(t_p - 1)} \quad (41)$$

$$FDE = \frac{\sum_{i=1}^N \left\| \vec{P}_p - \vec{P}_a \right\|_2}{N} \quad (42)$$

where  $\vec{P}_{ip}^t$  is the predicted trajectory vector at the time  $t$ ;  $\vec{P}_{ia}^t$  is the actual trajectory vector at the time  $t$ ;  $\vec{P}_p$  is the predicted trajectory vector in 2 seconds;  $\vec{P}_a$  is the actual trajectory vector in 2 seconds.

Fig. 8 shows a scene to predict the VRUs' trajectories based on the integrated model. As shown in Fig. 8 (a), four subject pedestrians and one subject electric cyclist are crossing the zebra crossing, and one dynamic vehicle is approaching. Fig. 8 (b) shows the comparison between the actual trajectories and predicted trajectories of the VRUs in 2 seconds, and it can be seen that a good agreement is obtained. In this scene, the ADE of the VRUs is 0.21m, and the FDE of the VRUs

TABLE IV  
ADE AND FDE OF THE PREDICTED TRAJECTORIES WITH 10 METHODS

| Models                             | ADE (m)     | FDE (m)     |
|------------------------------------|-------------|-------------|
| B-LSTM <sup>[22]</sup>             | 0.38        | 0.49        |
| MLP-social-GRU <sup>[24]</sup>     | 0.31        | 0.46        |
| SFM <sup>[27]</sup>                | 0.36        | 0.43        |
| SFM <sup>[29]</sup>                | 0.44        | 0.57        |
| Social GAN <sup>[37]</sup>         | 0.37        | 0.48        |
| Social-Transformer <sup>[38]</sup> | 0.31        | 0.41        |
| SIT <sup>[39]</sup>                | 0.28        | 0.40        |
| IA-GRU                             | 0.26        | 0.42        |
| ISFM                               | 0.28        | 0.37        |
| <b>Integrated Model</b>            | <b>0.23</b> | <b>0.32</b> |

is 0.27m. It indicates that the actual trajectories can be well reproduced based on the integrated model with the acceptable errors.

To validate the accuracy of the integrated model, a series of comparative experiments using the collected traffic data are presented. TABLE IV shows the ADE and FDE of the VRUs' predicted trajectories with 10 methods. It can be seen that the integrated model outperforms the existing methods. Compared with the B-LSTM [22], the ADE and FDE reduce by 39.5% and 34.7%, respectively. Compared with the MLP-social-GRU [24], the ADE and FDE reduce by 25.8% and 30.4%, respectively. Compared with the SFM [27], the ADE and FDE reduce by 36.1% and 25.6%, respectively. Compared with the SFM [29], the ADE and FDE reduce by 47.7% and 43.9%, respectively. Compared with the Social GAN [37], the ADE and FDE reduce by 37.8% and 33.3%, respectively. Compared with the Social-Transformer [38], the ADE and FDE reduce by 25.8% and 22.0%, respectively. Compared with the SIT [39], the ADE and FDE reduce by 17.9% and 20.0%, respectively. Compared with the IA-GRU, the ADE and FDE reduce by 11.5% and 23.8%, respectively. Compared with the ISFM, the ADE and FDE reduce by 17.9% and 13.5%, respectively. Overall, the integrated model improves the accuracy of the VRU trajectory prediction by at least 11%. As the data-driven integrated model considers VRU intention and VRU heterogeneity, and it also combines the advantages of the IA-GRU and ISFM, while other existing methods do not. Thus, more accurate results of the trajectory prediction are obtained. In addition, the computational load of the integrated model is 0.026 seconds, the computational load of the IA-GRU and ISFM are 0.021 seconds and 0.017 seconds, respectively. Though the computational load of the integrated model is higher than the IA-GRU and ISFM, considering the accuracy and run-time complexity, the integrated model can still satisfy the real-time of the autonomous vehicle. The operating system of the workstation used in this paper is the Ubuntu18.04, and the hardware is the GPU NVIDIA RTX 2080Ti workstation.

To further evaluate the accuracy of the proposed model, two public datasets (ETH and UCY) which are widely used in the existing methods are selected. Among them, ETH dataset was proposed by the relevant researchers of ETH Zurich, and UCY

TABLE V  
PREDICTION RESULTS OF THE BASELINE METHODS  
BASED ON TWO PUBLIC DATASETS

| Models                             | ADE (m)     | FDE (m)     |
|------------------------------------|-------------|-------------|
| B-LSTM <sup>[22]</sup>             | 0.64        | 1.25        |
| MLP-social-GRU <sup>[24]</sup>     | 0.59        | 1.08        |
| SFM <sup>[27]</sup>                | 0.69        | 1.12        |
| SFM <sup>[29]</sup>                | 0.78        | 1.47        |
| Social GAN <sup>[37]</sup>         | 0.59        | 1.18        |
| Social-Transformer <sup>[38]</sup> | 0.51        | 0.93        |
| SIT <sup>[39]</sup>                | 0.35        | 0.81        |
| IA-GRU                             | 0.34        | 0.86        |
| ISFM                               | 0.39        | 0.80        |
| <b>Integrated Model</b>            | <b>0.32</b> | <b>0.75</b> |

dataset was proposed by the relevant researchers of Cyprus University [23], [37]. In the two datasets, there are a total of 1536 pedestrians with different scenes, such as walking together, walking in the opposite direction and so on. These datasets just contain pedestrians, there are no cyclists, electric cyclists, or dynamic vehicles, so only pedestrian-pedestrian interaction is considered in the integrated model. As the frame rate is 0.4 seconds per frame, during the evaluation process, eight frames are observed for predicting the next twelve frames. In other words, a trajectory of 3.2 seconds is observed to predict the future trajectory of the next 4.8 seconds. TABLE V shows the prediction results of the baseline methods based on the two public datasets. Here the two public datasets were combined into one dataset (ETH+UCY) which included 1536 pedestrians' trajectories in total. It can be seen that the SFM [29] has the largest prediction errors, and the integrated model outperforms the existing methods in the ADE and FDE. Compared with the second-best performing SIT [39], the ADE and FDE reduce by 8.6% and 7.4%, respectively.

To prove the performance of the proposed model, an ablation experiment is conducted with the existing integrated methods. TABLE VI presents the comparative experiments with 4 integrated methods based on the collected traffic data. It can be seen that the proposed model achieves the highest accuracy of the predicted trajectories. Compared with the GBDT [40], the ADE and FDE reduce by 25.8% and 23.8%, respectively. Compared with the XGBOOST [41], the ADE and FDE reduce by 17.9% and 20.0%, respectively. Compared with the RF-LSTM [42], the ADE and FDE reduce by 32.4% and 30.4%, respectively. In fact, as the integrated model belongs to the heterogeneous integration algorithm, the output results of each base models (IA-GRU and ISFM) are quite different, and different features can be explored and fitted from different angles and directions. The integrated model can conduct the secondary generalization learning on the output results of the base models, which can reduce the deviation errors and improve the accuracy of trajectory prediction. Considering the advantages of the heterogeneous integration algorithm, therefore, better prediction results are obtained based on the integrated model.

TABLE VI  
ADE AND FDE OF VRUS' PREDICTED TRAJECTORIES  
WITH 4 INTEGRATED METHODS

| Models                  | ADE (m)     | FDE (m)     |
|-------------------------|-------------|-------------|
| GBDT <sup>[40]</sup>    | 0.31        | 0.42        |
| XGBOOST <sup>[41]</sup> | 0.28        | 0.40        |
| RF-LSTM <sup>[42]</sup> | 0.34        | 0.46        |
| <b>Integrated Model</b> | <b>0.23</b> | <b>0.32</b> |

TABLE VII  
ADE AND FDE OF PREDICTED TRAJECTORIES  
WITH VRU HETEROGENEITY

| VRUH | Models                  | ADE (m)     | FDE (m)     |
|------|-------------------------|-------------|-------------|
| No   | IA-GRU                  | 0.31        | 0.46        |
| No   | ISFM                    | 0.34        | 0.43        |
| No   | Integrated Model        | 0.28        | 0.36        |
| Yes  | IA-GRU                  | <b>0.26</b> | <b>0.42</b> |
| Yes  | ISFM                    | <b>0.28</b> | <b>0.37</b> |
| Yes  | <b>Integrated Model</b> | <b>0.23</b> | <b>0.32</b> |

TABLE VIII  
ADE AND FDE OF PREDICTED TRAJECTORIES  
WITH 4 DIFFERENT CROSSINGS

| Different crossings             | ADE (m) | FDE (m) |
|---------------------------------|---------|---------|
| Zebra Crossing on Bo Le Road    | 0.22    | 0.30    |
| Zebra Crossing on Qing He Road  | 0.25    | 0.33    |
| Intersection on Hui Wang Road   | 0.26    | 0.34    |
| Intersection on Huan Cheng Road | 0.28    | 0.35    |

To verify the effect of the VRU heterogeneity (age and gender) on the trajectory prediction, an ablation experiment is presented based on the traffic collected data. As shown in TABLE VII, VRUH represents the VRU heterogeneity, it can be seen that the proposed models (IA-GRU, ISFM, Integrated Model) considering the VRUH achieve more accurate prediction results than those do not consider. For the integrated model which considers the VRUH, compared with the integrated model which does not consider, the ADE and FDE improves by 17.9% and 11.1%, respectively. In other words, the VRUH is a key factor in improving the accuracy of the VRU trajectory prediction.

In this paper, to prove the good generalization of the proposed model, a serial of experiments with different crossings are designed. TABLE VIII presents the results of ADE and FDE with 4 different crossings. Here the two selected intersections are both without zebra crossings, and the forces from the zebra crossing boundary are removed. It is found that the integrated model can accurately predict the VRUs' trajectories with low errors, which can prove the scalability of the proposed model. In fact, the proposed model has been trained using an un-signalized zebra crossing data, and no matter which zebra crossings or intersections are selected, the interaction criterions between the VRU-VRU and VRU-vehicle are consistent. Therefore, the integrated model used

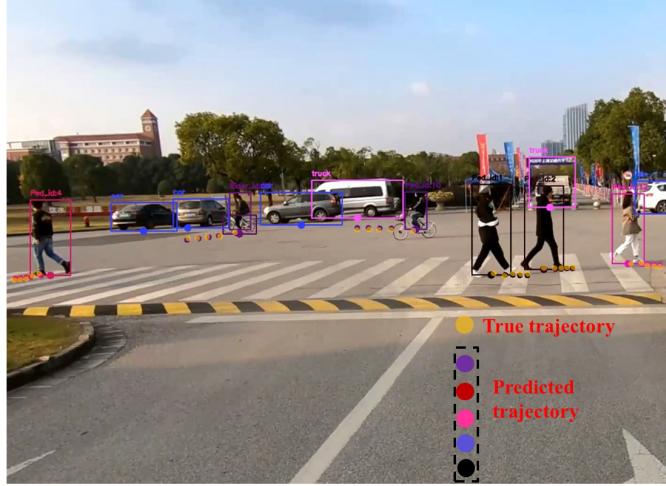


Fig. 9. VRU trajectory prediction based on the vision-based integrated model.

for the VRU trajectory prediction is effective when facing with different crossings.

To prove whether the integrated model can work in real-time, a real-time vision system which combines the target detection, multi-target tracking, VRU heterogeneity recognition and distance measurement is developed [43], [44], [45], [46]. The accuracy and prediction-time of the integrated model can be evaluated online. Fig. 9 shows the results of the VRU trajectory prediction based on the vision-based integrated model. The ADE is 0.27m, and the FDE is 0.38m. The average prediction-time based on the vision-based integrated model is 0.033 seconds, the average prediction-time based on the vision-based IA-GRU and ISFM are 0.028 seconds and 0.025 seconds, respectively. Considering the accuracy and real-time performance, the vision-based integrated model can meet the practical application of the autonomous vehicle. In other words, the vision-based integrated model can work in real-time for the VRU trajectory prediction.

## V. CONCLUSION

This paper systematically investigates the VRU trajectory prediction for the autonomous driving based on a data-driven integrated approach, and the conclusions are given below:

- 1) An IA-GRU used for the VRU trajectory prediction is developed, VRU intention, VRU heterogeneity, VRU-VRU interactions, VRU-dynamic vehicle interactions are taken into account. In addition, Attention is presented to obtain the influence weights of the above influencing factors.
- 2) A micro-dynamic ISFM used for the VRU trajectory prediction is developed, the influence of road users, zebra crossing, VRU heterogeneity are considered. The average desired speed and relaxation time of the VRUs by age type are presented, and the parameters of the ISFM are calibrated.
- 3) A data-driven integrated model based on the IA-GRU and ISFM is proposed for the VRU trajectory prediction. AdaBoost is presented to prevent the model from overfitting. The verified results indicate that the

TABLE IX  
MEANINGS OF SOME IMPORTANT NOTATIONS

| Notations                                     | Meanings   |
|---|--|
| $M_{i1} \dots M_{i5t}$                        | Multi-feature variables                                      |
| $O_{it}$                                      | Collection of VRU $i$ 's multi-feature variables             |
| $H_{it}$                                      | Intention encoding vector                                    |
| $P_{it}$                                      | Intent prediction probability distribution                   |
| $IN_{it}$                                     | VRU intention  |
| $X$   | All the VRUs' trajectories                                   |
| $V$   | All the dynamic vehicles' trajectories                       |
| $L$   | All the VRUs' age and gender type                            |
| $h_{ie}^t$                                    | Hidden state   |
| $H_{ie}$                                      | Collection of hidden state sequences of the VRU $i$          |
| $D_{ij}^t$                                    | Distance between the subject VRU $i$ and surrounding VRU $j$ |
| $v_{ij}^t$                                    | Relative velocity between the subject VRU $i$ and VRU $j$    |
| $D_{ivg}^t$                                   | Distance between the subject VRU $i$ and dynamic vehicle $g$ |
| $\theta_{ij}^t$                               | Collection of five features                                  |
| $a_t$   | Contribution weight of the features                          |
| $\vec{F}_{ic}$                                | Combined force   |
| $\vec{F}_d$                                   | Self-driving force from the destination                      |
| $\vec{F}_{ij}$                                | Repulsive force from the other conflicting VRUs              |
| $\vec{F}_{iv}$                                | Repulsive/accelerated force from the conflicting vehicle     |
| $\vec{F}_{iz}$                                | Repulsive/attractive force from the zebra crossing boundary  |
| $A_{ij}$ and $B_{ij}$                         | Strength coefficients of the other conflicting VRUs          |
| $A_{iv}^r; B_{iv}^r; A_{iv}^a$ and $B_{iv}^a$ | Strength coefficients of the conflicting vehicles            |
| $A_{iz}^r; B_{iz}^r; A_{iz}^a$ and $B_{iz}^a$ | Strength coefficients of the zebra crossing                  |
| $\theta_{VRU}$                                | Collection of the non-measurable parameters                  |
| $ADE$   | Average Displacement Error                                   |
| $FDE$   | Final Displacement Error                                     |

proposed model outperforms the existing models and the prediction accuracy is improved by at least 11% based on the collected traffic data, which is helpful for protecting the safety of the VRUs in the autonomous driving domain.

However, the proposed model still has some limitations. For example, the proposed model can only predict the trajectories of VRUs, and it cannot predict the trajectories of multiple road users (e.g., motorcycles and vehicles). In addition, the proposed model should also be used in the scene of signalized-zebra crossings, where there is clear priority for VRUs. The influence of signal and right-turning vehicles should be considered in the model. Potential future work will focus on improving the generalization/scalability of the proposed model.

## APPENDIX

In this paper, to make the mathematical notations easy to follow, the meanings of some important notations are described in TABLE IX.

## REFERENCES

- [1] Y. Yuan and J. Zhang, "A novel initiative braking system with non-degraded fallback level for ADAS and autonomous driving," *IEEE Trans. Ind. Electron.*, vol. 67, no. 6, pp. 4360–4370, Jun. 2020.
- [2] G. Xie, H. Gao, L. Qian, B. Huang, K. Li, and J. Wang, "Vehicle trajectory prediction by integrating physics- and maneuver-based approaches using interactive multiple models," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5999–6008, Jul. 2018.

- [3] X. Zhang, H. Chen, W. Yang, W. Jin, and W. Zhu, "Pedestrian path prediction for autonomous driving at un-signalized crosswalk using W/CDM and MSFM," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 5, pp. 3025–3037, May 2021.
- [4] X. Hu, B. Tang, L. Chen, S. Song, and X. Tong, "Learning a deep cascaded neural network for multiple motion commands prediction in autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 12, pp. 7585–7596, Dec. 2021.
- [5] Y. Xia, Z. Qu, Z. Sun, and Z. Li, "A human-like model to understand surrounding vehicles' lane changing intentions for autonomous driving," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4178–4189, May 2021, doi: 10.1109/TVT.2021.3073407.
- [6] Y. Deng, T. Zhang, G. Lou, X. Zheng, J. Jin, and Q.-L. Han, "Deep learning-based autonomous driving systems: A survey of attacks and defenses," *IEEE Trans. Ind. Informat.*, vol. 17, no. 12, pp. 7897–7912, Dec. 2021.
- [7] M. Goldhammer, S. Köhler, S. Zernetsch, K. Doll, B. Sick, and K. Dietmayer, "Intentions of vulnerable road users-detection and forecasting by means of machine learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 3035–3045, Jul. 2020.
- [8] J. B. Liu, X. Y. Mao, Y. Q. Fang, D. L. Zhu, and M. Q.-H. Meng, "A survey on deep-learning approaches for vehicle trajectory prediction in autonomous driving," in *Proc. IEEE Int. Conf. Robot. Biomim.*, Dec. 2021, pp. 978–985.
- [9] R. Quintero, I. Parra, D. F. Llorca, and M. A. Sotelo, "Pedestrian path prediction based on body language and action classification," in *Proc. IEEE 17th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 679–684.
- [10] Y. Huang, J. Du, Z. Yang, Z. Zhou, L. Zhang, and H. Chen, "A survey on trajectory-prediction methods for autonomous driving," *IEEE Trans. Intell. Vehicles*, vol. 7, no. 3, pp. 652–674, Sep. 2022.
- [11] F. Leon and M. Gavrilescu, "A review of tracking and trajectory prediction methods for autonomous driving," *Mathematics*, vol. 9, no. 6, pp. 660–696, Mar. 2021.
- [12] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, "Human motion trajectory prediction: A survey," *Int. J. Robot. Res.*, vol. 39, no. 8, pp. 895–935, Jun. 2020.
- [13] N. Schneider and D. M. Gavrila, "Pedestrian path prediction with recursive Bayesian filters: A comparative study," in *Proc. German Conf. Pattern Recognit.* Berlin, Germany: Springer, Sep. 2013, pp. 174–183.
- [14] C. Ning-Bo, Q. Zhao-Wei, C. Yong-Heng, Z. Li-Ying, S. Xian-Min, and B. Qiao-Wen, "Destination and route choice models for bidirectional pedestrian flow based on the social force model," *IET Intell. Transp. Syst.*, vol. 11, no. 9, pp. 537–545, Aug. 2017.
- [15] Z. Zhou, Y. Zhou, Z. Pu, and Y. Xu, "Simulation of pedestrian behavior during the flashing green signal using a modified social force model," *Transportmetrica A, Transp. Sci.*, vol. 15, no. 2, pp. 1019–1040, Nov. 2019.
- [16] R. Hug, S. Becker, W. Hubner, and M. Arens, "Particle-based pedestrian path prediction using LSTM-MDL models," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2684–2691.
- [17] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, arXiv:1412.3555.
- [18] B. Hussain, M. K. Afzal, S. Ahmad, and A. M. Mostafa, "Intelligent traffic flow prediction using optimized GRU model," *IEEE Access*, vol. 9, pp. 100736–100746, 2021.
- [19] Y. X. Ma, X. G. Zhu, S. B. Zhang, R. G. Yang, W. P. Wang, and D. Manocha, "TrafficPredict: Trajectory prediction for heterogeneous traffic-agents," *Comput. Sci.*, vol. 33, pp. 1–8, Apr. 2019.
- [20] X. Song et al., "Pedestrian trajectory prediction based on deep convolutional LSTM network," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3285–3302, Jun. 2021.
- [21] A. Bhattacharyya, M. Fritz, and B. Schiele, "Long-term on-board prediction of people in traffic scenes under uncertainty," *Comput. Sci.*, vol. 13, pp. 4194–4202, Jun. 2018.
- [22] K. Saleh, M. Hossny, and S. Nahavandi, "Contextual recurrent predictive model for long-term intent prediction of vulnerable road users," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 8, pp. 3398–3408, Aug. 2020.
- [23] S. Haddad and S.-K. Lam, "Self-growing spatial graph networks for pedestrian trajectory prediction," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 1140–1148.
- [24] Y. Zhang and L. Zheng, "Pedestrian trajectory prediction with MLP-social-GRU," in *Proc. 13th Int. Conf. Mach. Learn. Comput.*, Feb. 2021, pp. 368–372.
- [25] J. S. Zhu, S. Y. Chen, W. Tu, and K. Sun, "Tracking and simulating pedestrian movements at intersections using unmanned aerial vehicles," *Remote. Sens.*, vol. 925, pp. 1–19, Apr. 2019.
- [26] Y. Liu, W. Zhang, G. Cui, and Y. Zhang, "Multipedestrian online tracking based on social force-predicted deformable key-points mapping via compressive sensing," *IEEE Syst. J.*, vol. 15, no. 2, pp. 1905–1916, Jun. 2021.
- [27] D. Yang, Ü. Özguner, and K. Redmill, "A social force based pedestrian motion model considering multi-pedestrian interaction with a vehicle," *ACM Trans. Spatial Algorithms Syst.*, vol. 6, no. 2, pp. 1–27, Feb. 2020.
- [28] H. Chen, X. Zhang, W. Yang, W. Jin, W. Zhu, and B. Zhao, "W/CDM-MSFM-driven pedestrian path prediction at signalized crosswalk with mixed traffic flow," *Transportmetrica B, Transp. Dyn.*, vol. 9, no. 1, pp. 172–197, Oct. 2020.
- [29] W. Zeng, P. Chen, G. Yu, and Y. Wang, "Specification and calibration of a microscopic model for pedestrian dynamic simulation at signalized intersections: A hybrid approach," *Transp. Res. C, Emerg. Technol.*, vol. 80, pp. 37–70, Jul. 2017.
- [30] L. Huang, J. Wu, F. You, Z. Lv, and H. Song, "Cyclist social force model at unsignalized intersections with heterogeneous traffic," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 782–792, Apr. 2017.
- [31] S. Marisamynathan and P. Vedagiri, "Modeling pedestrian delay at signalized intersection crosswalks under mixed traffic condition," *Proc. Social Behav. Sci.*, vol. 104, pp. 708–717, Dec. 2013.
- [32] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
- [33] H. Q. Tran and C. Ha, "Reducing the burden of data collection in a fingerprinting-based VLP system using a hybrid of improved co-training semi-supervised regression and adaptive boosting algorithms," *J. Opt. Commun.*, vol. 488, Jun. 2021, Art. no. 126857.
- [34] B. Xiong, R. Li, D. Ren, H. Liu, T. Xu, and Y. Huang, "Prediction of flooding in the downstream of the three Gorges reservoir based on a back propagation neural network optimized using the AdaBoost algorithm," *Natural Hazards*, vol. 107, no. 2, pp. 1559–1575, Mar. 2021.
- [35] J. Yang et al., "No-reference quality assessment for screen content images using visual edge model and AdaBoosting neural network," *IEEE Trans. Image Process.*, vol. 30, pp. 6801–6814, 2021.
- [36] Y. Li, X.-Y. Lu, J. Wang, and K. Li, "Pedestrian trajectory prediction combining probabilistic reasoning and sequence learning," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 3, pp. 461–474, Sep. 2020.
- [37] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, "Social GAN: Socially acceptable trajectories with generative adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 2255–2264.
- [38] H. B. Sun and F. C. Sun, "Social-transformer: Pedestrian trajectory prediction in autonomous driving scenes," in *Proc. Int. Conf. Cogn. Syst. Sign. Process.*, Jan. 2022, pp. 177–190.
- [39] T. Su, Y. Meng, and Y. Xu, "Pedestrian trajectory prediction via spatial interaction transformer network," in *Proc. IEEE Intell. Veh. Symp.*, Jul. 2021, pp. 154–159.
- [40] W. Zhang, J. Yu, A. Zhao, and X. Zhou, "Predictive model of cooling load for ice storage air-conditioning system by using GBDT," *Energy Rep.*, vol. 7, pp. 1588–1597, Nov. 2021.
- [41] P. Liu and W. Fan, "Extreme gradient boosting (XGBoost) model for vehicle trajectory prediction in connected and autonomous vehicle environment," *Promet Traffic Transp.*, vol. 33, no. 5, pp. 767–774, Oct. 2021.
- [42] D. Choi, J. Yim, M. J. Baek, and S. S. Lee, "Machine learning-based vehicle trajectory prediction using V2V communications and on-board sensors," *Electron.*, vol. 10, no. 4, pp. 420–438, Feb. 2021.
- [43] A. Bochkovskiy, C. Wang, and H. Liao, "YOLOv4: Optimal speed and accuracy of object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Apr. 2020, pp. 1–17.
- [44] A. Bewley, Z. Y. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and real time tracking," *IEEE Image. Process.*, pp. 3464–3468, Feb. 2016.
- [45] D. Li, X. Chen, and K. Huang, "Multi-attribute learning for pedestrian attribute recognition in surveillance scenarios," in *Proc. ACPR*, Nov. 2015, pp. 111–115.
- [46] Q. Li, "Research on monocular vision-based realtime distance measurement algorithm," M.S. dissertation Eng., Harbin Inst. Technol., Harbin, China, Dec. 2014.



**Hao Chen** received the M.S. degree in mechanical engineering and automation from Jiangsu University, Zhenjiang, China, in 2011, and the Ph.D. degree from the School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China, in 2021.

He is currently an Assistant Professor with the School of Mechanical Engineering, University of Shanghai for Science and Technology, China. His research interests include vulnerable road user trajectory prediction for intelligent vehicles, vulnerable road user modeling and simulation, local trajectory planning, and intelligent decision technology.



**Chuan Hu** received the B.E. degree in automotive engineering from Tsinghua University, Beijing, China, in 2010, the M.E. degree in vehicle operation engineering from the China Academy of Railway Sciences, Beijing, in 2013, and the Ph.D. degree in mechanical engineering from McMaster University, Hamilton, ON, Canada, in 2017.

He was a Post-Doctoral Fellow with the Department of Systems Design Engineering, University of Waterloo, from July 2017 to July 2018, and The University of Texas at Austin from August 2018 to June 2020. He was an Assistant Professor with the Department of Mechanical Engineering, University of Alaska Fairbanks, from June 2020 to June 2022. He is currently an Associate Professor with the School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China. His research interests include the perception, decision-making, path planning, and motion control of connected and automated vehicles (CAVs), autonomous driving, eco-driving, human-machine trust and cooperation, shared control, and machine learning applications in CAVs. He has published more than 60 papers in these research areas. He is also an Associate Editor of IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON TRANSPORTATION ELECTRIFICATION, and IEEE TRANSACTIONS ON INTELLIGENT VEHICLES.



**Yinhua Liu** received the B.S. degree in industrial design from Jilin University, Changchun, China, in 2005, and the Ph.D. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2013.

She is currently an Associate Professor with the School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai. Her studies are sponsored by the National Natural Science Foundation of China, the Natural Science Foundation of Shanghai, and SAIC General Motors Corporation. Her research interests include engineering-driven data analytics in manufacturing, trajectory and motion planning of robotics and other agents, and big data analytics for process design and quality improvement, especially in the area of automobile industry.



**Xi Zhang** (Senior Member, IEEE) received the B.Sc. degree in applied mathematics, the B.E. degree in information and control engineering, and the M.E. and Ph.D. degrees in power electronics and electric power drive from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2002, 2004, and 2007.

From September 2007 to July 2009, he held a post-doctoral position with the Department of Electrical and Computer Engineering, University of Michigan–Dearborn, Dearborn, MI, USA. He is currently a Full Professor with the Institute of Intelligent and Connected Automobile and the National Engineering Laboratory for Automotive Electronics and Control Technology, SJTU. His research interests include intelligent and connected automobile, intelligent transportation, power management strategies, power electronics devices, and electric motor control systems for alternative-fuel vehicles.