

## Article

# Pedestrian Trajectory Prediction Based on an Intention Randomness Influence Strategy

Yingjian Deng <sup>1</sup>, Li Zhang <sup>2,\*</sup> , Jie Chen <sup>1,3</sup>, Yu Deng <sup>1</sup> and Jing Liu <sup>4</sup>

<sup>1</sup> The Information Materials and Intelligent Sensing Laboratory of Anhui Province, Anhui University, Hefei 230601, China

<sup>2</sup> School of Integrated Circuits, Anhui University, Hefei 230601, China

<sup>3</sup> The 38th Research Institute of China Electronics Technology Group Corporation, Hefei 230601, China

<sup>4</sup> The Shandong Branch of the National Computer Network Emergency Technology Coordination Center, Jinan 250004, China

\* Correspondence: 04068@ahu.edu.cn

**Abstract:** Pedestrian trajectory prediction is a key technical prerequisite for autonomous vehicle trajectory planning. However, a pedestrian is a changeable individual, and their intentions exhibit certain degrees of randomness and uncertainty, which leads to the issue that modeling only past trajectories does not enable the effective description of the random intentions and future trajectory directions of the pedestrian. Therefore, this paper proposes a flexible and embeddable stochastic intention vector construction strategy for modeling sudden pedestrian intention changes in real scenes and for better fitting the stochastic properties of pedestrian behaviors. First, we dynamically fuse historical trajectory information with random factors and construct an intention change probability based on the historical trajectory fitting errors of pedestrians, aiming to explicitly model the associated direction and velocity changes caused by random pedestrian intentions. Second, a new intention loss function is designed to guide the model to adaptively learn the probability of intention changes, which is used to dynamically describe pedestrian intention changes. Our proposed method is generalizable and can be applied as an embeddable module to any baseline pedestrian trajectory prediction method. The experimental results obtained on multiple large-scale public pedestrian trajectory prediction datasets demonstrate that our strategy achieves consistent performance improvements over different baselines.

**Keywords:** pedestrian trajectory prediction; pedestrian intention; random intention vector; RNN; GNN



**Citation:** Deng, Y.; Zhang, L.; Chen, J.; Deng, Y.; Liu, J. Pedestrian Trajectory Prediction Based on an Intention Randomness Influence Strategy. *Electronics* **2024**, *13*, 2008. <https://doi.org/10.3390/electronics13112008>

Academic Editor: Prasan Kumar Sahoo

Received: 16 April 2024

Revised: 9 May 2024

Accepted: 20 May 2024

Published: 21 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

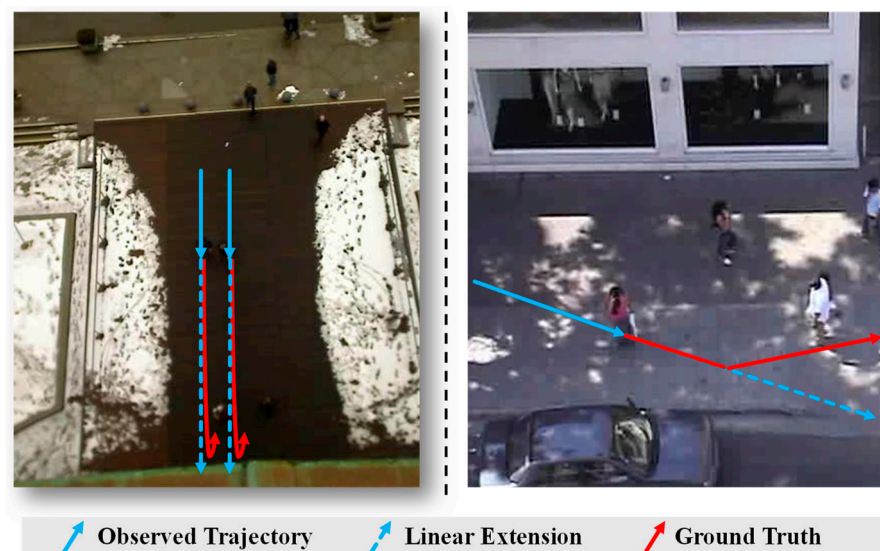
## 1. Introduction

The accurate prediction of pedestrians' future trajectories through their historical trajectories is an urgent problem to be solved with respect to autonomous driving technology and has attracted a large amount of attention from researchers. In recent years, the field of pedestrian trajectory prediction has also garnered considerable attention and new methods have constantly emerged.

The current methods are generally based on recurrent neural networks (RNNs); they model pedestrians' historical trajectories to simulate the hidden states of pedestrian movements, and then, based on an attention module, they weight the interactions between pedestrians [1–3] or directly model pedestrian nodes based on different graph neural networks. The adjacency matrix of the interactions between pedestrians is calculated and used to predict the future trajectories of pedestrians [4–6]. Although some of the current methods have achieved good results, because a pedestrian is an uncertain individual, their movements are random. Specifically, in a normal traffic scene, the intention to change at any time due to the different semantic constraints of a scene instructs pedestrians to choose

to slow down or change their walking directions, directly leading to sudden changes in their future trajectories.

At present, research methods based on modeling the motion patterns of pedestrians according only to their historical motion trajectories may not be suitable for pedestrian trajectory prediction tasks in actual scenarios. This is because the existing methods do not consider the random intentions of pedestrians themselves. Specifically, the existing methods do not model the influence of the random intentions of pedestrians on their future trajectories. As shown in Figure 1, the information provided by a historical trajectory alone cannot be used to cope with trajectory changes caused by sudden pedestrian intention changes. The linear extension in the figure is the continuation of the pedestrian's current intention. The pedestrian's intention clearly changes during the subsequent walking process, and their future trajectory changes accordingly. Notably, this random change in intent cannot be predicted solely from the pedestrian's historical trajectory information. Therefore, the method of predicting future trajectories based on historical trajectory data is ideal for prediction scenarios where pedestrian intentions do not change. Once the intentions of pedestrians change, information such as the directions and speeds provided by historical trajectories no longer has absolute reference significance. Therefore, this randomness should be fully considered in both autonomous driving applications and intelligent robot applications.



**Figure 1.** Background of our approach. The figure shows the impact of a sudden pedestrian intention change in an actual scene. The solid blue line represents the observed historical track, the dashed blue line represents the linear extension of the current track, and the solid red line represents the subsequent real track.

To our knowledge, the current methods rarely consider the impacts of pedestrian intention changes on their directions and speeds [3,5,7]. On the one hand, the randomness of pedestrian movement is difficult to describe; on the other hand, pedestrian movements have certain limitations under social rules, so the current methods pay more attention to the interactions between pedestrians and their physical environment. However, this does not mean that the randomness of pedestrians themselves can be ignored. In complex scenes, an automatic driving system must have extremely high driving safety, especially in cases where pedestrians interact with the system. Therefore, pedestrian trajectory prediction systems for automatic driving applications must consider the influence of random pedestrian intentions to ensure safety.

To this end, this paper proposes a stochastic intention vector modeling strategy for characterizing the randomness of pedestrian intentions. Specifically, we generate an intention vector that simulates pedestrian intention changes by combining the displacement

information of the last frame of the historical trajectory with the random normal distribution and then predicting the subsequent pedestrian trajectory. The movement direction and speed changes caused by the intention vector are introduced to simulate the pedestrian intention changes. Then, we designed a module that models the probability of pedestrian intention changes. First, the direct manifestation of a pedestrian intention change is often reflected on their trajectory; that is, trajectory performance is not linear. Pedestrians with variable historical trajectories are more likely to change their future intentions. Therefore, the linear fitting error of the historical trajectory can be used as the probability of a pedestrian intention change to determine whether to use the intention vector. Finally, the impacts of sudden changes in pedestrians' directions and speeds can be continuously reduced through the use of the intention vector during training. Our proposed method is a general plug-in module that is designed to model the influence of the randomness of pedestrian intentions on pedestrian trajectory prediction, and our method can be embedded into different existing methods. In this paper, we also implement some applications based on the state-of-the-art baseline modeling method. The experimental results show that our method demonstrates good improvements over the baseline model.

The main contributions of this paper are as follows.

1. We propose a random intention vector construction strategy that combines historical trajectory information with random factors, where the aim is to predict a trajectory by directly considering the direction and speed changes caused by the changes in a pedestrian's own intentions.
2. We designed an intention change probability based on the fitting error induced by the historical trajectory of the pedestrian and introduce an intention loss function to self-update the intention change probability, ensuring that the process of modeling the trajectory change caused by the randomness of pedestrian movement is targeted.
3. The proposed intention vector construction method can be embedded into all trajectory prediction benchmark methods as a plug-in module to supplement the pedestrian randomness modeling process to improve the pedestrian trajectory prediction accuracies of the benchmark methods.

## 2. Related Work

### 2.1. Pedestrian Trajectory Prediction

Pedestrian trajectory prediction is based mainly on the combination of a pedestrian's own movement pattern and their interactions with the outside world. In the early methods, pedestrian movements were mainly simulated by artificial motion functions [8,9]. These predictions were usually deterministic and did not involve multiple modes, so it was difficult to simulate the randomness of pedestrian movements. Moreover, these methods did not consider the historical trajectory information of pedestrians but rather focused on the continuation of the current motion state.

With the rapid development of deep learning, pedestrian trajectory prediction methods have also been improved. Social long short-term memory (LSTM) [10] utilizes RNNs to model the motion state of each pedestrian to obtain the hidden state representing their motion mode. Then, the proposed social pooling module combines the hidden states of adjacent pedestrians within a certain spatial distance to simulate interactions between the pedestrians. The social-generative adversarial network (GAN) [11], SoPhie [1] and Social-Ways [2] all retain the part of Social LSTM that uses RNNs to determine pedestrian motion states; in addition, they use GANs as their overall network architectures. The pooling module proposed for Social-GAN comprehensively considers global pedestrian information, which is different from the local pooling operation of Social LSTM, so the generated trajectory is more in line with social norms. On the basis of pedestrians' historical trajectories, SoPhie also introduces real scene information, integrating social interaction attention and physical scene attention. Social-Ways simulates pedestrian interactions and introduces an attention mechanism to comprehensively consider the distances between pedestrians, the angles between the motion directions of pedestrians and the shortest

distances between pedestrians to better understand the interactions between pedestrians. SR-LSTM [12] employs a data-driven LSTM state refinement network to enable the utilization of the current intentions of neighbors through a message-passing framework. Other methods adopt the variational autoencoder (VAE) framework in combination with RNNs to directly predict tracks [7,13]. Additionally, some methodologies incorporate convolutional neural networks to model the spatial characteristics of pedestrians. In [14], a convolutional LSTM network was designed to extract the temporal and multichannel spatial features of pedestrians to predict their future trajectories. In [15], a two-dimensional convolutional model was directly introduced instead of an RNN, and data normalization and data enhancement techniques were combined to predict the trajectories of pedestrians.

Some methods have also begun to consider the characteristics of graph networks. Pedestrians and their influences on other pedestrians can be embedded as the nodes and edges of graph networks, which can better simulate the interactions between people. Social-BiGAT [16] and STGAT [3] still use RNNs as their main frameworks and employ only a graph attention network (GAT) to address pedestrian interactions in trajectory prediction tasks, while Social-STGCNN [4] directly models pedestrian trajectories as graphs and performs trajectory prediction on graph structures. In the prediction stage, a multimodal prediction strategy based on a Gaussian distribution is used to replace the confirmed trajectory predicted by the model to improve the resulting prediction accuracy. When a graph convolutional network (GCN) is used to model pedestrian interactions according to the distances between them, the impacts between pedestrians are the same, leading to redundant interactions. SGCN [5] improved upon Social-STGCNN by utilizing sparse GCNs to learn trajectory representations.

## 2.2. Intention Prediction

Some relevant studies have focused on the uncertainty encountered in driving scenes, investigated drivers' behavioral characteristics, and identified drivers' intentions [17,18]. Correctly predicting a pedestrian's intentions is also crucial for achieving autonomous driving, as it provides a safe way to plan vehicle routes. Pedestrian intentions can be used to better judge the subsequent destinations and paths of pedestrians. Therefore, it is extremely important to predict pedestrian intentions by combining historical trajectory information and even physical scenes.

Some early works used hidden Markov models to predict pedestrian intentions [19,20]. Recently, several methods for use with intelligent driving assistance systems have emerged; they mainly predict pedestrian intentions regarding whether to cross a road [21,22]. Other works aimed at predicting the intentions of pedestrians with respect to their future destinations; such methods pay more attention to scene information, and the final target locations of pedestrians are considered before performing trajectory prediction [23,24]. However, these methods use the historical trajectory and scene information of a pedestrian to judge their current intention information. This strategy often ignores the fact that the pedestrian's own movement is stochastic, i.e., the pedestrian's current intention may change in subsequent movements. Such random intention fluctuations influence the pedestrian's subsequent movements and future destinations, thereby further compromising the predictive efficacy of the baseline trajectory prediction model. Consequently, the random intention module contained in our approach is designed to model the impacts of sudden pedestrian intention changes. It can be used as a plug-and-play module to complement the ability of the baseline method to model intention randomness.

## 3. Methods

In this section, we present the process for constructing a random intention vector and the full implementation details of the baseline model. Some of the main notations used in this paper are summarized in Table 1.

**Table 1.** Summary of the main notations.

Notation	Description
$X_n$	Past trajectory of agent $n$
$E_n$	Interaction information of agent $n$
$F(\cdot)$	Trajectory prediction network
$Y_n$	Future trajectory of agent $n$
$N$	Total number of agents in a scene
$\vec{R}_{obs}$	A coordinate component of the displacement of frame 8
$\vec{\eta}$	A vector of random numbers sampled from a standard normal distribution
$w, b$	Learnable weights and bias in a Linear model
$\mu, v$	Learnable parameters in a custom Linear* model
$\vec{I}$	A random intention vector that combines coordinate components and random number factors
$\vec{p}$	Initialized fitting error vector with values between 0 and 1
$\vec{P}_I$	A learnable vector that represents the probability of each pedestrian's intention to change
$\hat{P}_I$	The true value of the intention change probability obtained from the fitting error of the future trajectory
$\vec{I}_s$	The output of the sparse intention module—a sparse random intention vector
$\varphi_n, \theta_n$	Random number between 0 and 1
$\vec{P}_I(n)$	The probability of the intention of agent $n$ changing
$\vec{I}_{out}$	The final output of the module obtained by splicing $\vec{I}_s$ with zero vectors
$\vec{R}$	The displacement vector obtained by encoding the historical trajectory of a pedestrian
$\vec{R}_I$	Displacement vectors that superimpose intention changes

### 3.1. Problem Definition

Pedestrian track prediction operates based mainly on historical track information and physical scene information to predict subsequent position coordinates. In the current method, the historical time frames  $t_1 \in \{1, 2, \dots, t_{obs}\}$  are generally given, and for each person  $n \in [1, \dots, N]$  included in the scene of interest, the spatial coordinates  $(x_{t_1}^n, y_{t_1}^n)$  in time frame  $t_1$  are defined as historical track coordinates. Based on these coordinates, the coordinate information in the subsequent time frame  $t_2 \in \{t_{obs} + 1, t_{obs} + 2, \dots, t_{pred}\}$  is predicted.

Pedestrian interactions  $E_n$  can be obtained from the input pedestrian trajectory or scene information. After passing through the trajectory prediction network  $F_\theta(\cdot)$  in sequence, the future trajectory of a pedestrian can be modeled as follows:

$$Y_n = F_\theta(X_n, E_n) \quad (1)$$

where  $X_n = \{(x_{t_1}^n, y_{t_1}^n) \in R \mid t_1 = 1, 2, \dots, t_{obs}\}$  represents the historical trajectory of the  $n$ th pedestrian in the scene and  $Y_n = \{(x_{t_2}^n, y_{t_2}^n) \in R \mid t_2 = t_{obs} + 1, t_{obs} + 2, \dots, t_{pred}\}$  represents the predicted future trajectory of the  $n$ th pedestrian. The current method basically uses relative positions, namely, displacement values, to replace absolute positions. The pedestrian movement pattern obtained according to the given historical track information predicts the subsequent displacement on the basis of the direction and speed provided by the displacement of the last frame, enabling the acquisition of future position coordinates.

### 3.2. Construction of a Random Intention Vector and Intention Change Probabilities

This section describes how to construct a random intention vector based on a given historical trajectory to simulate the impacts of pedestrians' actual intention changes. The intuitive expression of a pedestrian intention change encountered in a real scene represents

the fact that the current speed and direction of a pedestrian change, so we need to simulate such changes manually.

The displacement of the last frame in the historical trajectory reflects the direction and magnitude of the current speed of the pedestrian, that is, the direct manifestation of the pedestrian’s current intention. In real scenes, the change in a pedestrian’s intention is intuitively represented by their trajectory change. With a physical background, as long as the coordinate component of the pedestrian speed changes, it will produce changes in the magnitude and direction of the subsequent velocity.

The specific structure of our random intention module is shown in Figure 2. First, the displacement information of the last frame of the input historical trajectory is combined with a random number factor generated from the standard normal distribution, and then the combined information is passed through a biased linear layer to obtain a random intention vector  $\vec{I}$ . In the fit error module, we initialize the intentional change probability of each pedestrian based on the magnitudes of the second-order fitness error values of all the coordinate points in the input history trajectory and then run it through a linear layer with a random bias term to obtain the intentional change probability  $\vec{P}_I$ . In the sparse intention module, the need for the random intention vector  $\vec{I}$  is determined based on the magnitude of the intention change probability  $\vec{P}_I$  for each pedestrian. The output  $\vec{I}_s$  of the sparse intention module is eventually concatenated with a zero vector to maintain the same length as that of the historical trajectory.

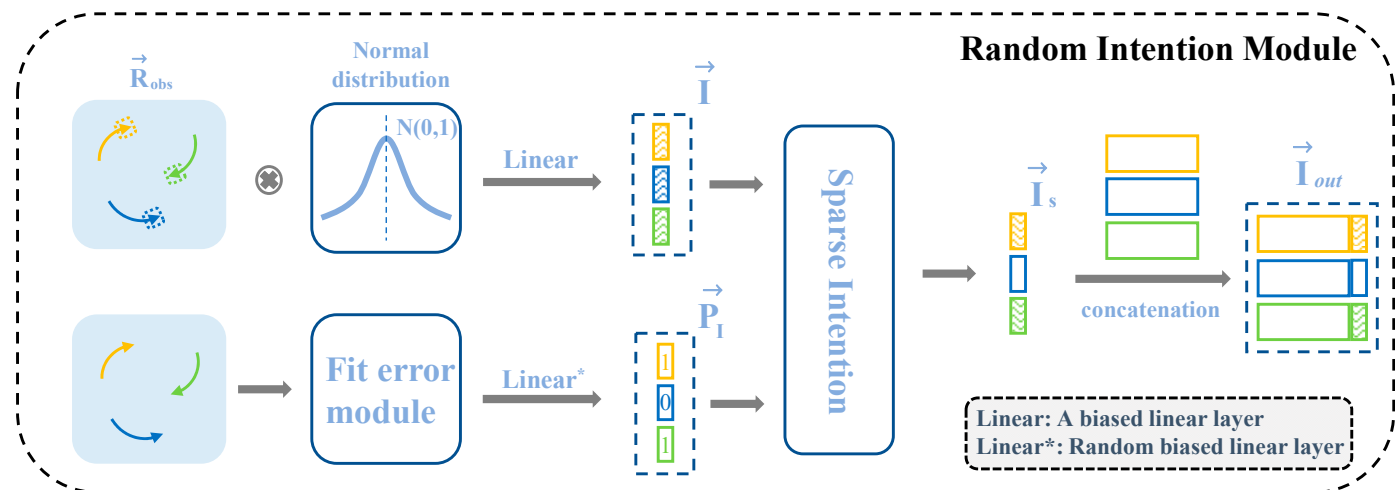


Figure 2. The random intention module.

### 3.2.1. Constructing the Random Intention Vector $\vec{I}$

As shown in Figure 2, we first obtain the displacement in the last frame of the historical trajectory through the historical trajectory coordinates, then use the coordinate component  $\vec{R}_{obs}$  of the displacement in the last frame as the reference information, and finally combine  $\vec{R}_{obs}$  with random number factors to simulate a pedestrian speed change. With a social background, most pedestrian tracks are basically unchanged or change slightly, while a few pedestrians exhibit significant changes in their tracks. At the same time, the pedestrian trajectory change simulation process needs to be able to perform forward and reverse operations on the original state, so the generated random number factors should preferably include positive and negative distributions. In our method, the standard normal distribution is used to generate each random number factor  $\vec{\eta}$ , and the effects of different data distributions are experimentally evaluated in Section 4.2.

We sample a random number factor  $\vec{\eta}$  from the standard normal distribution, feed  $\vec{R}_{obs}$  into a biased linear layer, and multiply the result by  $\vec{\eta}$  to obtain the random intention vector:

$$\vec{I} = \left( w\vec{R}_{obs} + b \right) \otimes \vec{\eta} \quad (2)$$

where  $\otimes$  represents the multiplication of the corresponding elements of two vectors;  $w$  and  $b$  represent the learnable weight parameters and bias terms in the linear layer, respectively; and  $\vec{R}_{obs}$  contains the speed and direction of the pedestrian in the last frame, that is, the intention information. Combined with the randomness caused by the random number factor  $\vec{\eta}$ , the random intention vector combined with the current pedestrian state can be initially obtained. Afterward, the values of  $w$  and  $b$  are constantly updated during the model training process to enable the model to self-adjust to the randomness in the random intention vector and avoid extremely unreasonable intention changes.

### 3.2.2. Constructing the Intention Change Probability $\vec{P}_I$

Next, a standard is set based on the historical trajectory data to define which pedestrians may be more inclined to change their intentions, making our random intention vector more targeted and thus improving the randomness of pedestrians without creating too many redundant influences. Since a pedestrian's trajectory is a direct reflection of their intention, we design a fitting error module. First, we perform second-order fitting for the coordinate points of each pedestrian's historical trajectory based on the least-squares method to obtain the fitting error and then initialize the fitting error to a vector  $\vec{p}$  between 0 and 1 according to the size of the error. Then, the vector  $\vec{p}$  is passed through a linear layer, Linear\*, with a random bias for obtaining the vector representing the intention change probability of each pedestrian as follows:

$$\vec{P}_I = \nu \vec{p} + \mu \varphi_n \quad (3)$$

where  $\nu$  and  $\mu$  are learnable parameters in the defined randomly biased linear layer (Linear\*), and  $\varphi_n$  is a set of random numbers drawn from a uniform distribution with the interval  $[0, 1)$ . The weight  $\nu$  and random bias term  $\mu \varphi_n$  are introduced here mainly to make the model more flexible for updating the probability of pedestrian intention changes during the training process according to different real scenarios. At the same time, the linear layer uses the artificially designed random bias term  $\mu \varphi_n$ , which also makes it possible to use for pedestrians with small probabilities.

### 3.3. Intention Loss Function

Notably, the loss function of the baseline model cannot effectively update a pedestrian's intention change probability vector  $\vec{P}_I$  on its own. Therefore, to help the model capture the relationship between the historical trajectory offset and the future trajectory offset of a pedestrian, we introduce a new intention loss to ensure that the model can more flexibly update the pedestrian's intention change probability  $\vec{P}_I$  during the training process. The loss is calculated as follows:

$$L_{Intention} = \left\| \vec{P}_I - \hat{P}_I \right\|_2 \quad (4)$$

where  $\hat{P}_I$  represents the true value of the intention change probability obtained according to the fitting error between the last frame of the pedestrian's historical trajectory and the future trajectory. The final loss function is defined as a linear combination of the loss functions  $L_{Baseline}$  and  $L_{Intention}$  of the baseline model, namely,  $L = L_{Baseline} + \lambda L_{Intention}$ .

We set a penalty factor  $\lambda = 0.1$  to balance the primary and secondary relationships between the two losses.

### 3.4. Sparse Intention

As shown in Figure 2, the obtained random intention vector  $\vec{I}$  and the intention change probability vector  $\vec{P}_I$  are also passed through a sparse intention module to obtain a sparse random intention vector  $\vec{I}_s$ . The designed sparse module determines whether the corresponding pedestrian should use the generated random intention vector or use the zero vector without any influence according to each probability contained in  $\vec{P}_I$ . In other words, element-by-element threshold judgments are needed for  $\vec{P}_I$ . Instead of setting a specific threshold, a random number  $\theta_n$  between 0 and 1 is generated for each pedestrian  $n$ . In this way, when  $\vec{P}(n) \geq \theta_n$ , the  $n$ th element of  $\vec{I}_s$  is set to the  $n$ th element of  $\vec{I}$ ; otherwise, it is set to 0. That is,

$$\vec{I}_s(n) = \begin{cases} \vec{I}(n), & \vec{P}_I(n) \geq \theta_n \\ 0, & \vec{P}_I(n) < \theta_n \end{cases} \quad (5)$$

Finally, a zero vector is generated and spliced with the sparse random intention vector  $\vec{I}_s$  to obtain the module output  $\vec{I}_{out}$ , which is mainly used to subsequently superimpose it with the complete historical pedestrian trajectory displacement.

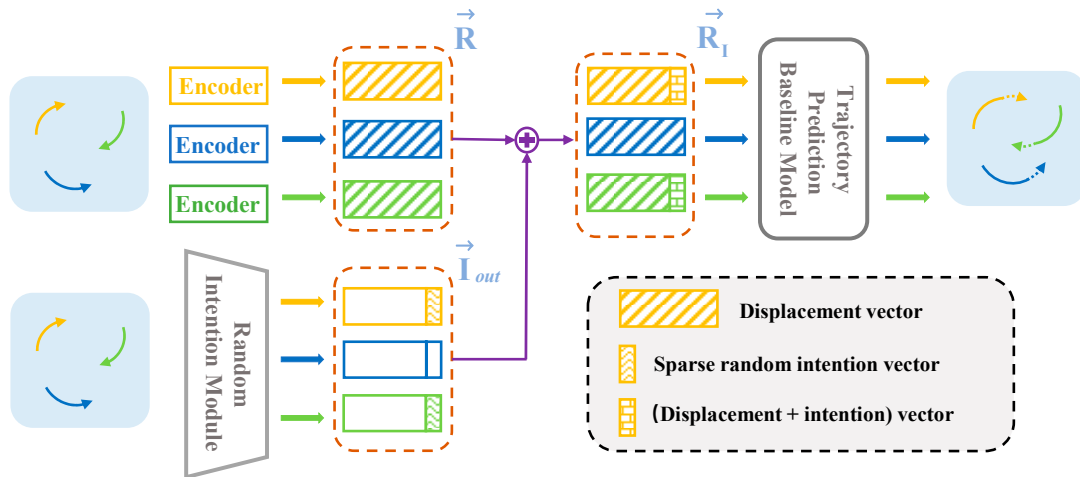
### 3.5. Implementation Details

To evaluate the generality of our method for constructing random intention vectors, we add the proposed method as a module to the existing method for evaluation purposes. As shown in Figure 3, first, the historical trajectory coordinates of pedestrians need to be coded as a displacement vector  $\vec{R}$ , and the displacement is defined as the coordinate difference between two adjacent frames to facilitate consistency with the input form of the baseline model data during the test. At the same time, the historical trajectory coordinates of pedestrians are input into the random intention module to obtain the intention module output  $\vec{I}_{out}$  spliced with the zero vector. Afterward, the displacement vector  $\vec{R}$  and the intention module output  $\vec{I}_{out}$  are superimposed to obtain  $\vec{R}_I$ , namely:

$$\vec{R}_I = \vec{R} \oplus \vec{I}_{out} \quad (6)$$

where  $\oplus$  represents the elementwise addition operation. In this way, the intention vector can induce pedestrian speed and direction changes, thus simulating the influences of pedestrian intention changes.

We then input the displacement vector  $\vec{R}_I$ , which adds the intention changes to the baseline model we want to evaluate. Here, we add  $\vec{R}_I$  to three advanced baseline methods—the GNN-based SGCN [5] and the RNN-based STGAT [3] and SVAE [7] methods—for training purposes to obtain the predicted trajectories. Additionally, when evaluating our methods, we use the same hyperparameters as those of the baseline models for training, employing a fixed random number seed, allowing our results to be efficiently reproduced, thus providing a fair comparison between the results of the baseline models and those produced after the addition of our intention module.



**Figure 3.** Practical use of our intention vector construction method. The historical trajectory coordinates of a pedestrian are encoded to obtain a displacement vector. Additionally, the trajectory coordinates of the pedestrian are sent to the random intention module and superimposed on the displacement vector of the pedestrian.

#### 4. Experiments

In this section, we introduce comprehensive evaluation experiments, which are performed on the proposed intention vector construction method to confirm its effectiveness and generality through a detailed comparison. In Section 4.1, we first briefly introduce the employed pedestrian trajectory prediction datasets (ETH [25] and UCY [26]) and the current mainstream evaluation indicators. In Section 4.2, the comparison results obtained by adding our intention module to the baseline model and the original baseline model are shown in detail. The displacement coordinate components of the last frame of the input historical trajectory ( $\vec{R}_{obs}$ ) obtained using different forms are experimentally evaluated. The impacts of the random number factors  $\eta$  generated by different distributions are evaluated. At the same time, we remove some components from the random intention module to obtain different variants and evaluate the contributions of different components to the final attained performance. Finally, in Section 4.3, we present the visualization results obtained on the ETH and UCY datasets.

##### 4.1. Datasets and Evaluation Metrics

**Datasets:** To train and evaluate our method, we conducted experiments on two public datasets, ETH and UCY, which were sampled every 0.4 s from pedestrian trajectories produced in different real scenarios. The ETH dataset includes two real ETH and Hotel scenarios. The UCY dataset includes ZARA1, ZARA2, and UNIV, which are three real scenes. To evaluate our method, we adopted the mainstream training idea. For the five real scenario datasets, we trained and validated the model by evaluating it on four of the datasets and tested it on the remaining dataset. In the experiment, the historical trajectory lasted for 3.2 s (8 frames), and the predicted future trajectory was 4.8 s (12 frames).

**Evaluation Metrics:** We used the same evaluation criteria as those of the baseline method to evaluate the prediction results, namely, the mean displacement error (ADE) [27] and the final displacement error (FDE) [10], where the ADE calculates the mean Euclidean distance between all the predicted trajectory coordinates and all the true trajectory coordinates. The final displacement distance is calculated by the Euclidean distance between the predicted endpoint coordinates and the real destination coordinates. The specific definitions of these metrics are as follows:

$$ADE = \frac{\sum_{n \in N} \sum_{t=t_{obs}}^{t_{pred}} \|\hat{Y}_n^t - Y_n^t\|_2}{N * (t_{pred} - t_{obs})} \tag{7}$$

$$\text{FDE} = \frac{\sum_{n \in N} \left\| \hat{Y}_n^{t=t_{pred}} - Y_n^{t=t_{pred}} \right\|_2}{N} \quad (8)$$

At the same time, the assessment strategy of the baseline method was also adopted during the test; that is, 20 samples were generated, and the samples closest to the real trajectory were selected for evaluation purposes.

#### 4.2. Quantitative Evaluation

In this section, we present all the experiments we performed to determine the effectiveness of the proposed random intention module.

Experimental evaluation of the intention vector construction process: To evaluate the effectiveness of our approach, as shown in Table 2, we conducted experiments on the official source codes of three representative CNN-based and RNN-based baseline models (SGCN, STGAT, and SVAE) and compared the results of the baseline models with those of Intention-SGCN, Intention-STGAT, and Intention-SVAE after adding our approach. An \* in Table 2 indicates the results we reproduced using the official code of the method. Due to the different settings of some hyperparameters or the use of different experimental equipment, our directly reproduced results may have been different from the results produced in the original papers. Therefore, to ensure the fairness of the comparison and to better reflect the actual results of our methods, all the hyperparameters in our evaluations of Intention-SGCN and Intention-STGAT were consistent with those in the baseline models. For different baseline models, our intention vector construction method improved the error metrics produced during model testing. In particular, Intention-SGCN improved the average ADE/FDE values for the five real-world scenarios by 0.02/0.12, respectively, over those of SGCN \*. Compared with that of STGAT \*, the FDEs of Intention-STGAT improved by 0.03/0.05 over those of STGAT\*, and that of Intention-SVAE improved by 0.01 over that of SVAE \*. The Stanford UAV dataset [28], which contains eight different real-world scenarios, was also used as a pedestrian trajectory prediction benchmark [7,13,29]. For the sake of the comprehensiveness of the evaluation, we kept the same dataset segmentation settings as those in [29] and evaluated the trajectory prediction effectiveness of Intention-SVAE with the addition of our intention module on the SDD dataset, using the latest SVAE approach as the baseline model. As shown in Table 3, in the evaluation conducted on the SDD dataset, our Intention-SVAE achieved an effective FDE improvement of 0.02 over the baseline SVAE model. In the evaluation of our method, the y-coordinate component of the displacement in the last frame of the historical trajectory was used to construct the intention vector as the benchmark information  $\vec{R}_{obs}$ . Table 4 presents our attempts to adopt different forms of displacement coordinate components for the last frame of the historical trajectory ( $\vec{R}_{obs}$ ) and provides the relevant experimental evaluation results.

Evaluation of the results produced when selecting different  $\vec{R}_{obs}$ . As shown in Table 4, before determining the final scheme, we tried to use different forms of the coordinate components in the last frame of the historical trajectory as benchmark information for constructing intention vectors. The experiments included three schemes: Intention-SGCN-X, Intention-SGCN-Y, and Intention-SGCN-XY. Intention-SGCN-X involved the use of the y-coordinate of the displacement of frame 8 as  $\vec{R}_{obs}$ , Intention-SGCN-Y means used the y-coordinate as  $\vec{R}_{obs}$ , and Intention-SGCN-XY used both the x- and y-coordinates as  $\vec{R}_{obs}$ . It can be seen that any of the schemes could yield performance improvements over the baseline model and improvements in the randomness of pedestrian intention changes. These results also demonstrate the universality of our intention vector construction method. Overall, the better results produced by Intention-SGCN-Y may have been obtained because pedestrian movements are always regular in specific scenarios, such as pedestrians walking along the street when shopping and pedestrians walking parallel to the zebra crossing, that is, pedestrians walking along a “road”. When the given dataset is labeled, under the set coordinate system, the forward direction for most pedestrians is consistent with

the direction of the x-coordinate system, so the x-component of pedestrian displacement mainly affects the speed, while changes in the y-component can not only affect the speeds of pedestrians but also cause more direction changes. At the same time, although pedestrian movements are random, the walking speeds of pedestrians are relatively low, and their speed changes are continuous. Even if a change occurs, it is more represented by a shift in direction. Therefore, Intention-SGCN-Y is more suitable for actual situations; that is, the selected  $\vec{R}_{obs}$  must be set according to the vertical direction of the route in an actual application scene.

Evaluation of the  $\vec{\eta}$  generated from different data distributions: To determine the final data distribution we adopted in Section 3.2 and further verify the validity of our conjecture regarding the standard normal distribution, as shown in Table 5, we designed relevant experiments to evaluate the actual effects of the random number factors  $\vec{\eta}$  generated by different data distributions, including the uniform distribution, random normal distribution, and standard normal distribution, on the resulting models. The standard normal distribution achieved the best comprehensive effects in different real scenes, which was consistent with our conjecture. The standard normal distribution provided random positive and negative factors, and small random factors were obtained with high probability so that a large proportion of the pedestrian trajectories were only slightly changed, which was in line with the pedestrian movement patterns observed in real scenes.

**Table 2.** Evaluation results produced by several advanced baseline models. Our two-stage optimization strategy improved the predictive efficacy levels of all three baseline methods. Lower values of both the ADE and FDE metrics signify better results. The table highlights the best results for ADE/FDE in bold.

Baseline 1	Performance (ADE/FDE)					
	ETH	HOTEL	ZARA1	ZARA2	UNIV	AVG
Social LSTM [10]	1.09/2.35	0.86/1.91	0.41/0.88	0.52/1.11	0.61/1.31	0.70/1.52
SGAN [11]	0.87/1.62	0.67/1.37	0.35/0.68	0.42/0.84	0.76/1.52	0.61/1.21
SoPhie [1]	0.70/1.43	0.76/1.67	0.30/0.63	0.38/0.78	0.54/1.24	0.54/1.15
CGNS [30]	0.62/1.40	0.70/0.93	0.32/0.59	0.35/0.71	0.48/1.22	0.49/0.97
STSGN [31]	0.75/1.63	0.63/1.01	0.30/0.65	0.26/0.57	0.48/1.08	0.48/0.99
PITF [32]	0.73/1.65	0.30/0.59	0.38/0.81	0.31/0.68	0.60/1.27	0.46/1.00
Social-BiGAT [16]	0.69/1.29	0.49/1.01	0.30/0.62	0.36/0.75	0.55/1.32	0.48/1.00
STAR [33]	<b>0.56/1.11</b>	<b>0.26/0.50</b>	0.41/0.90	0.31/0.71	0.52/1.15	0.41/0.87
STGCNN [4]	0.64/1.11	0.49/0.85	0.34/0.53	0.30/0.48	0.44/0.79	0.44/0.75
Conv2D-tobs-NR-Ks5 [15]	<b>0.56/1.11</b>	<b>0.24/0.46</b>	0.46/0.99	0.35/0.75	0.58/1.23	0.44/0.91
SGCN [5]	0.63/1.03	0.32/0.55	0.29/0.53	0.25/0.45	<b>0.37/0.70</b>	0.37/0.65
SGCN *	0.71/1.34	0.29/0.54	0.27/0.48	0.24/0.43	0.38/0.70	0.38/0.70
Intention-SGCN	0.64/ <b>0.83</b>	0.28/ <b>0.46</b>	<b>0.27/0.47</b>	<b>0.24/0.42</b>	0.38/0.73	<b>0.36/0.58</b>
Baseline 2	Performance (ADE/FDE)					
	ETH	HOTEL	ZARA1	ZARA2	UNIV	AVG
STGAT [3]	<b>0.65/1.12</b>	<b>0.35/0.66</b>	0.34/0.69	<b>0.29/0.60</b>	<b>0.52/1.10</b>	<b>0.43/0.83</b>
STGAT *	0.80/1.42	0.37/0.70	0.33/0.66	0.29/0.61	0.55/1.17	0.47/0.91
Intention-STGAT	0.68/1.18	0.37/0.70	<b>0.32/0.66</b>	0.30/0.61	0.54/1.14	0.44/0.86
Baseline 3	Performance (ADE/FDE)					
	ETH	HOTEL	ZARA1	ZARA2	UNIV	AVG
PECNet [34]	0.54/0.87	0.18/0.24	0.22/0.39	0.17/0.30	0.35/0.60	0.29/0.48
Trajectron++ [13]	0.54/0.94	0.16/0.28	0.21/0.42	0.16/0.31	0.28/0.55	0.27/0.50
Social-Implicit [29]	0.66/1.44	0.20/0.36	0.25/0.50	0.22/0.43	0.31/0.60	0.33/0.67
SVAE [7]	<b>0.49/0.77</b>	<b>0.15/0.24</b>	<b>0.19/0.37</b>	<b>0.15/0.28</b>	<b>0.25/0.47</b>	<b>0.25/0.43</b>
SVAE *	0.50/0.85	<b>0.15/0.23</b>	0.20/ <b>0.37</b>	0.16/0.29	0.25/0.48	<b>0.25/0.44</b>
Intention-SVAE	0.50/ <b>0.75</b>	0.16/ <b>0.23</b>	0.20/ <b>0.37</b>	<b>0.15/0.29</b>	0.26/0.49	<b>0.25/0.43</b>

**Table 3.** Results of evaluation experiments conducted on the SDD dataset, where \* denotes the results reproduced using the baseline model.

Method	Social-Ways [2]	DAG-Net [35]	Social-Implicit [29]	SVAE *	Intention-SVAE
ADE	0.62	0.53	0.47	0.33	0.33
FDE	1.16	1.04	0.89	0.53	0.51

**Table 4.** The SGCN was used as the baseline model to evaluate the results obtained when using different forms of the displacement components of frame 8 to construct the intention vector.

	Performance (ADE/FDE)					
	ETH	HOTEL	ZARA1	ZARA2	UNIV	AVG
SGCN * (baseline)	0.71/1.34	0.29/0.54	<b>0.27/0.48</b>	<b>0.24/0.43</b>	<b>0.38/0.70</b>	0.38/0.70
Intention-SGCN-X	0.71/1.23	0.30/0.53	<b>0.27/0.50</b>	<b>0.24/0.43</b>	<b>0.38/0.74</b>	0.38/0.68
Intention-SGCN-XY	0.66/1.18	<b>0.28/0.50</b>	0.28/0.50	<b>0.24/0.44</b>	0.39/0.73	0.37/0.67
Intention-SGCN-Y (ours)	<b>0.64/0.83</b>	<b>0.28/0.46</b>	<b>0.27/0.47</b>	<b>0.24/0.42</b>	<b>0.38/0.73</b>	<b>0.36/0.58</b>

**Table 5.** The SGCN was used as the baseline model to evaluate the influences of the random number factors generated by different data distributions on the experimental results.

	Performance (ADE/FDE)					
	ETH	HOTEL	ZARA1	ZARA2	UNIV	AVG
SGCN * (baseline)	0.71/1.34	0.29/0.54	<b>0.27/0.48</b>	0.24/0.43	0.38/0.70	0.38/0.70
Intention-SGCN1	0.70/1.37	0.28/0.46	<b>0.27/0.50</b>	<b>0.23/0.41</b>	0.39/ <b>0.67</b>	0.37/0.68
Intention-SGCN2	0.65/1.21	<b>0.26/0.43</b>	<b>0.27/0.48</b>	<b>0.23/0.42</b>	<b>0.37/0.69</b>	<b>0.36/0.65</b>
Intention-SGCN3 (ours)	<b>0.64/0.83</b>	0.28/0.46	<b>0.27/0.47</b>	0.24/0.42	0.38/0.73	<b>0.36/0.58</b>

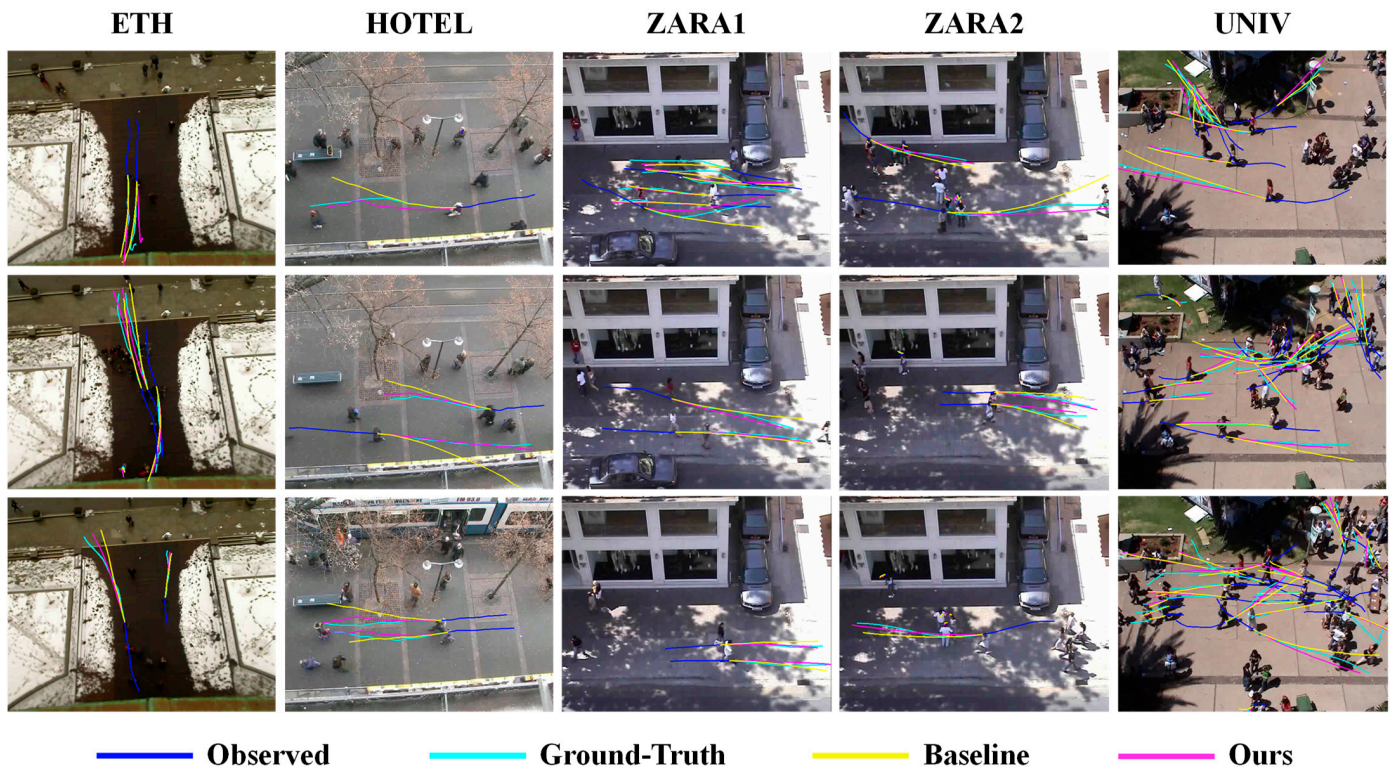
Evaluating the contributions of different components in the proposed module: As shown in Table 6, we evaluated three different variants of our random intention module, where (1) Intention-SGCN w/o RBLI indicates that the random biased linear layer (Linear\*) was removed, making it difficult for pedestrians with small initial intention change probabilities to change their intentions; (2) Intention-SGCN w/o RNF means that the random number factor  $\vec{\eta}$  was not used, so the random intention vector  $\vec{I}$  lost its randomness; and (3) Intention-SGCN w/o POI means that intentions were not used to directly change the probabilities, and each pedestrian had a random intention vector so that the constructed random intention vector was no longer targeted, resulting in some redundant effects. As shown by the results of our ablation experiments (presented in Table 6), removing any component from the model resulted in a substantial performance degradation. In particular, Intention-SGCN w/o POI did not conform to the pedestrian movement pattern in the real scene because each pedestrian considered their intention change, which directly led to a decline in the prediction performance achieved in scenarios other than the ETH scenario relative to that of the baseline model. This result shows that setting the intention change probability is crucial to the universality of the random intention module.

**Table 6.** Ablation experiments with variants of random intention modules with different components removed to evaluate the effect of different components in the module.

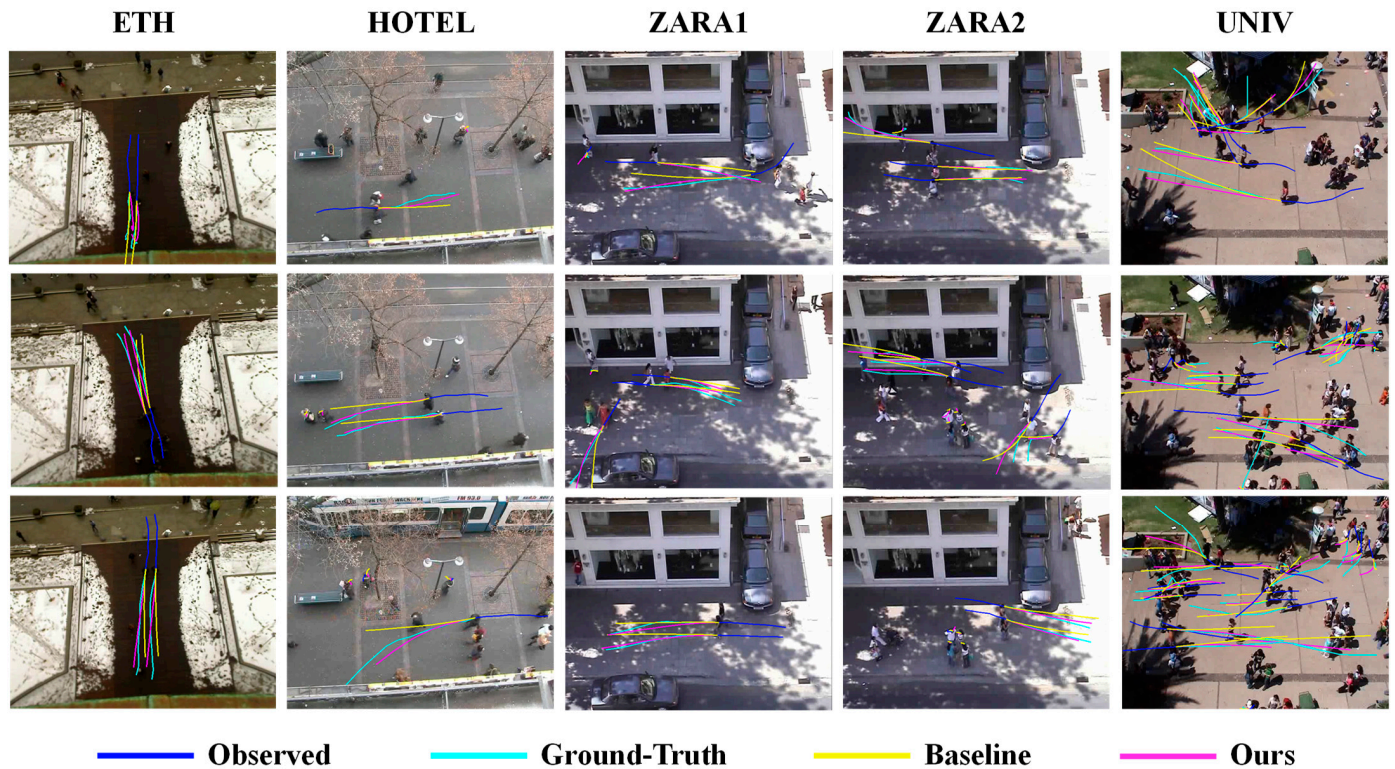
	Performance (ADE/FDE)					
	ETH	HOTEL	ZARA1	ZARA2	UNIV	AVG
SGCN * (baseline)	0.71/1.34	0.29/0.54	<b>0.27/0.48</b>	<b>0.24/0.43</b>	<b>0.38/0.70</b>	0.38/0.70
Intention-SGCN w/o RBLL	0.65/1.10	0.30/0.53	<b>0.27/0.49</b>	<b>0.24/0.43</b>	<b>0.38/0.72</b>	0.37/0.65
Intention-SGCN w/o RNF	0.66/1.18	0.32/0.56	<b>0.27/0.49</b>	<b>0.24/0.43</b>	0.41/0.75	0.38/0.68
Intention-SGCN w/o POI	<b>0.64/0.92</b>	0.33/0.60	<b>0.27/0.50</b>	<b>0.25/0.45</b>	0.39/0.74	0.37/0.64
Intention-SGCN (ours)	<b>0.64/0.83</b>	<b>0.28/0.46</b>	<b>0.27/0.47</b>	<b>0.24/0.42</b>	<b>0.38/0.73</b>	<b>0.36/0.58</b>

### 4.3. Visualization

We qualitatively visualize the results to show that our intention vector can indeed improve the consideration of the randomness of pedestrian intention changes. As shown in Figures 4 and 5, we visualize the trajectories predicted by Intention-SGCN and Intention-STGAT along with those of their corresponding baseline models in different scenarios. As illustrated in Figures 4 and 5, the utilization of our proposed methodology enables the predicted trajectories of the baseline model to be more closely aligned with the actual future trajectories of pedestrians when the true future trajectories exhibit significant deviations from the historical trajectories. Our method can effectively lessen the negative impact caused by the randomness of pedestrian movements and improve the overall accuracy of trajectory prediction.



**Figure 4.** Visual presentation of the trajectories produced by our Intention-SGCN method in different real-world scenarios. We compare the predicted trajectories generated by the Intention-SGCN and baseline SGCN methods on five datasets: ETH, HOTEL, ZARA1, ZARA2, and UNIV. These comparison results qualitatively show that our method can reduce the impact of random pedestrian intentions.



**Figure 5.** Visual presentation of the trajectories produced by our Intention-STGAT method in different real-world scenarios. We contrast the predicted trajectories generated by the Intention-STGAT and baseline STGAT methods on five datasets: ETH, HOTEL, ZARA1, ZARA2, and UNIV.

## 5. Conclusions

In this paper, we propose an intention vector construction method to model trajectory alterations caused by the randomness of pedestrian movements. This result is achieved through the designed random intention module, in which the intention change probabilities of pedestrians are obtained based on their historical trajectory information. At the same time, a pedestrian trajectory change is modeled according to the displacement information of the last frame of the historical trajectory and is superimposed on the displacement of the pedestrian through the sparse intention module. Our method is a plug-and-play module that can effectively improve the pedestrian trajectory prediction effects of different models. The experimental results show that our method achieves consistent performance improvements over the baseline models and is effective in different real-world scenarios. In addition, ablation experiments involving different intention vector construction methods verify the generalizability of our method. Similarly, the practical effectiveness of the method proposed in this paper as a plug-in module is constrained by the trajectory prediction efficacy of the baseline model. In future research, it would be beneficial to improve the process used to model scene information to gain a deeper understanding of the various possibilities of pedestrian movements.

**Author Contributions:** Investigation, Y.D. (Yingjian Deng), L.Z. and J.C.; methodology, Y.D. (Yingjian Deng); validation, Y.D. (Yingjian Deng); writing—original draft, Y.D. (Yingjian Deng); writing—review and editing, L.Z., J.C., Y.D. (Yu Deng) and J.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the National Natural Science Foundation of China under grant 62001003 and in part by the China Postdoctoral Science Foundation under grant 2020M671851 and in part by the National Science Foundation of China under grant U23B2007 and in part by the Key Research and Development Plan (Industry) Project of Yancheng under grant BE2023002.

**Data Availability Statement:** The data are available upon request from the authors.

**Conflicts of Interest:** Author Jie Chen was employed by the company The 38th Research Institute of China Electronics Technology Group Corporation, Author Jing Liu was employed by the company The Shandong Branch of the National Computer Network Emergency Technology Coordination Center. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Sadeghian, A.; Kosaraju, V.; Sadeghian, A.; Hirose, N.; Rezatofighi, H.; Savarese, S. Sophie: An attentive gan for predicting paths compliant to social and physical constraints. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seoul, Republic of Korea, 27 October–2 November 2019.
2. Amirian, J.; Hayet, J.B.; Pettré, J. Social ways: Learning multi-modal distributions of pedestrian trajectories with gans. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019.
3. Huang, Y.; Bi, H.; Li, Z.; Mao, T.; Wang, Z. STGAT: Modeling Spatial-Temporal Interactions for Human Trajectory Prediction. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 6271–6280.
4. Mohamed, A.; Qian, K.; Elhoseiny, M.; Claudel, C. Social-stgcnn: A social spatio-temporal graph convolutional neural network for human trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020.
5. Shi, L.; Wang, L.; Long, C.; Zhou, S.; Zhou, M.; Niu, Z.; Hua, G. SGCN: Sparse graph convolution network for pedestrian trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Montreal, QC, Canada, 10–17 October 2021.
6. Sun, J.; Jiang, Q.; Lu, C. Recursive social behavior graph for trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 660–669.
7. Xu, P.; Hayet, J.B.; Karamouzas, I. Socialvae: Human trajectory prediction using timewise latents. In *Computer Vision—ECCV 2022, Proceedings of the European Conference, Tel Aviv, Israel, 23–27 October 2022*; Springer Nature: Cham, Switzerland, 2022.
8. Mehran, R.; Oyama, A.; Shah, M. Abnormal crowd behavior detection using social force model. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, 20–25 June 2009.
9. Yamaguchi, K.; Berg, A.C.; Ortiz, L.E.; Berg, T.L. Who are you with and where are you going? In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20–25 June 2011.
10. Alahi, A.; Goel, K.; Ramanathan, V.; Robicquet, A.; Li, F.-F.; Savarese, S. Social LSTM: Human Trajectory Prediction in Crowded Spaces. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 961–971.
11. Gupta, A.; Johnson, J.; Li, F.-F.; Savarese, S.; Alahi, A. Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2255–2264.
12. Zhang, P.; Xue, J.; Zhang, P.; Zheng, N.; Ouyang, W. Social-aware pedestrian trajectory prediction via states refinement LSTM. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *44*, 2742–2759. [[CrossRef](#)] [[PubMed](#)]
13. Salzmann, T.; Ivanovic, B.; Chakraborty, P.; Pavone, M. Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data. In *Computer Vision—ECCV 2020, Proceedings of the 16th European Conference, Glasgow, UK, 23–28 August 2020*; Proceedings, Part XVIII 16; Springer: Cham, Switzerland, 2020; pp. 683–700.
14. Song, X.; Chen, K.; Li, X.; Sun, J.; Hou, B.; Cui, Y.; Zhang, B.; Xiong, G.; Wang, Z. Pedestrian trajectory prediction based on deep convolutional LSTM network. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 3285–3302. [[CrossRef](#)]
15. Zamboni, S.; Kefato, Z.T.; Girdzijauskas, S.; Norén, C.; Dal Col, L. Pedestrian trajectory prediction with convolutional neural networks. *Pattern Recognit.* **2022**, *121*, 108252. [[CrossRef](#)]
16. Kosaraju, V.; Sadeghian, A.; Martín-Martín, R.; Reid, I.; Rezatofighi, H.; Savarese, S. Social-bigat: Multimodal trajectory forecasting using bicycle-gan and graph attention networks. In Proceedings of the Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, Vancouver, BC, Canada, 8–14 December 2019.
17. Xin, L.; Wang, P.; Chan, C.Y.; Chen, J.; Li, S.E.; Cheng, B. Intention-aware long horizon trajectory prediction of surrounding vehicles using dual LSTM networks. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1441–1446.
18. Chen, J.; Tang, C.; Xin, L.; Li, S.E.; Tomizuka, M. Continuous decision making for on-road autonomous driving under uncertain and interactive environments. In Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1651–1658.
19. Kelley, R.; Tavakkoli, A.; King, C.; Nicolescu, M.; Nicolescu, M.; Bebis, G. Understanding human intentions via hidden markov models in autonomous mobile robots. In Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction, Amsterdam, The Netherlands, 12–15 March 2008; pp. 367–374.

20. Vasquez, D.; Fraichard, T.; Laugier, C. Incremental learning of statistical motion patterns with growing hidden markov models. *IEEE Trans. Intell. Transp. Syst.* **2009**, *10*, 403–416. [[CrossRef](#)]
21. Keller, C.G.; Gavrilu, D.M. Will the pedestrian cross? a study on pedestrian path prediction. *IEEE Trans. Intell. Transp. Syst.* **2013**, *15*, 494–506. [[CrossRef](#)]
22. Schneemann, F.; Heinemann, P. Context-based detection of pedestrian crossing intention for autonomous driving in urban environments. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Republic of Korea, 9–14 October 2016*; IEEE: Piscataway, NJ, USA, 2016; pp. 2243–2248.
23. Dendorfer, P.; Osep, A.; Leal-Taixé, L. Goal-gan: Multimodal trajectory prediction based on goal position estimation. In *Proceedings of the Asian Conference on Computer Vision, Kyoto, Japan, 30 November–4 December 2020*.
24. Dendorfer, P.; Elflein, S.; Leal-Taixé, L. Mg-gan: A multi-generator model preventing out-of-distribution samples in pedestrian trajectory prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision, Motreal, BC, Canada, 11–17 October 2021*; pp. 13158–13167.
25. Pellegrini, S.; Ess, A.; Schindler, K.; van Gool, L. You’ll never walk alone: Modeling social behavior for multi-target tracking. In *Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009*; pp. 261–268.
26. Lerner, A.; Chrysanthou, Y.; Lischinski, D. Crowds by Example. *Comput. Graph. Forum* **2007**, *26*, 655–664. [[CrossRef](#)]
27. Raksincharoensak, P.; Hasegawa, T.; Nagai, M. Motion planning and control of autonomous driving intelligence system based on risk potential optimization framework. *Int. J. Automot. Eng.* **2016**, *7*, 53–60. [[CrossRef](#)] [[PubMed](#)]
28. Robicquet, A.; Sadeghian, A.; Alahi, A.; Savarese, S. Learning social etiquette: Human trajectory understanding in crowded scenes. In *Computer Vision—ECCV 2016, Proceedings of the 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016*; Proceedings, Part VIII 14; Springer International Publishing: Cham, Switzerland, 2016.
29. Mohamed, A.; Zhu, D.; Vu, W.; Elhoseiny, M.; Claudel, C. Social-implicit: Rethinking trajectory prediction evaluation and the effectiveness of implicit maximum likelihood estimation. In *Computer Vision—ECCV 2022, Proceedings of the European Conference, Tel Aviv, Israel, 23–27 October 2022*; Springer Nature: Cham, Switzerland, 2022.
30. Li, J.; Ma, H.; Tomizuka, M. Conditional generative neural system for probabilistic trajectory prediction. In *Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macao, China, 3–8 November 2019*; IEEE: Piscataway, NJ, USA, 2019; pp. 6150–6156.
31. Zhang, L.; She, Q.; Guo, P. Stochastic trajectory prediction with social graph network. *arXiv* **2019**, arXiv:abs/1907.10233.
32. Liang, J.; Jiang, L.; Niebles, J.C.; Hauptmann, A.G.; Li, F.-F. Peeking into the future: Predicting future person activities and locations in videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seoul, Republic of Korea, 27 October–2 November 2019*.
33. Yu, C.; Ma, X.; Ren, J.; Zhao, H.; Yi, S. Spatio-temporal graph transformer networks for pedestrian trajectory prediction. In *Computer Vision—ECCV 2020, Proceedings of the 16th European Conference, Glasgow, UK, 23–28 August 2020*; Proceedings, Part XII 16; Springer International Publishing: Cham, Switzerland, 2020.
34. Mangalam, K.; Girase, H.; Agarwal, S.; Lee, K.H.; Adeli, E.; Malik, J.; Gaidon, A. It is not the journey but the destination: Endpoint conditioned trajectory prediction. In *Computer Vision—ECCV 2020, Proceedings of the 16th European Conference, Glasgow, UK, 23–28 August 2020*; Proceedings, Part II 16; Springer International Publishing: Cham, Switzerland, 2020.
35. Monti, A.; Bertugli, A.; Calderara, S.; Cucchiara, R. Dag-net: Double attentive graph neural network for trajectory forecasting. In *Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021*; IEEE: Piscataway, NJ, USA, 2021; pp. 2551–2558.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.