Heliyon 10 (2024) e30659

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Review article

Review of the application of neural network approaches in pedestrian dynamics studies

Shenshi Huang [a], Ruichao Wei [b,*], Liping Lian [a], Siuming Lo [c], Shouxiang Lu [d]

[a] School of Architectural Engineering, Shenzhen Polytechnic, Shenzhen, Guangdong, China
[b] School of Automobile and Transportation, Shenzhen Polytechnic, Shenzhen, Guangdong, China
[c] Department of Architecture and Civil Engineering, City University of Hong Kong, Kowloon, Hong Kong
[d] State Key Laboratory of Fire Science, University of Science and Technology of China, Hefei, China

ARTICLE INFO

Keywords:
Neural network
Pedestrian dynamics
Trajectory prediction
Pedestrian behavior

ABSTRACT

In recent years, artificial intelligence methods have been widely used in the study of pedestrian dynamics and crowd evacuation. Different neural network models have been proposed and tested using publicly available pedestrian datasets. These studies have shown that different neural network models present large performance differences for different crowd scenarios. To help future research select more appropriate models, this article presents a review of the application of neural network methods in pedestrian dynamics studies. The studies are classified into two categories: pedestrian trajectory prediction and pedestrian behavior prediction. Both categories are discussed in detail from a conceptual perspective, as well as from the viewpoints of methodology, measurement, and results. The review found that the mainstream method of pedestrian trajectory prediction is currently the LSTM-based method, which has adequate accuracy for short-term predictions. Furthermore, the deep neural network is the most popular method for pedestrian behavior prediction. This method can emulate the decision-making process in a complex environment, and it has the potential to revolutionize the study of pedestrian dynamics. Overall, it is found that new methods and datasets are still required to systemize the study of pedestrian dynamics and eventually ensure its wide-scale application in industry.

1. Introduction

Pedestrian dynamics is one of the most important research directions in traffic systems, and there are many approaches to study this issue. A traditional approach is to observe the surveillance video that can count the flow rate and movement characteristics of pedestrian flow in the different urban areas. By conducting the controlled experiments, the crowd characteristics of specific pedestrian scenarios can be studied in depth [1–3]. In addition, many modelling methods have been proposed to reproduce the pedestrian movement, such as, social force model [4,5], cellular automaton model [6–8], and agent-based model [9]. For the emergency situations, few scholars have conducted case studies [10,11], evacuation simulation [12–14], and crowd risk assessment [15]. The previous studies focused on the patterns and mechanisms of crowd movement, exploring the causes of group behaviors, and attempted to reproduce the phenomena through numerical modelling. The insights into dynamic mechanisms of pedestrians can facilitate the design of building structures and facilities. Furthermore, it is important to make online predictions for crowd movement for the daily and

* Corresponding author.
E-mail address: rcwei@mail.ustc.edu.cn (R. Wei).

https://doi.org/10.1016/j.heliyon.2024.e30659
Received 5 April 2023; Received in revised form 14 April 2024; Accepted 1 May 2024
Available online 5 May 2024
2405-8440/

probable emergency situations, to ensure the safe and efficient pedestrian flow capability in urban areas. Artificial neural network method is a common prediction approach that can provide a fast response to the problems. In recent years, many pedestrian-related topics have emerged in the field of computer science, such as pedestrian detection and trajectory analysis, and an increasing number of scholars have started to conduct the research on pedestrian dynamics and safety.

Neural network is a well-known machine learning algorithm. By learning a large number of labelled data samples of a specific problem, the neural network is trained and able to provide prompt timely answers to the problem. Previous studies have shown that neural networks can provide adequate predictive capabilities for non-linear systems. Compared to the traditional modelling methods for pedestrian prediction, the neural network approach requires an extremely short computation time, making it particularly suitable for applications in the fields of urban pedestrian flow management and edge computing. In addition, many online prediction methods for pedestrian problems have been proposed based on the neural network approach. Problems relevant to pedestrian behaviors are usually solved using deep neural networks (DNNs). DNNs are capable of solving the problems of pedestrian decision-making and safety assessment at a given moment [16–21]. Recurrent Neural Networks (RNNs) are suitable for solving problems over a continuous period of time. Based on the RNN concept, various improved LSTM approaches have been proposed and widely applied for pedestrian trajectory prediction [22–25]. In addition, convolutional neural networks (CNNs) are capable of handling problems related to pedestrian images, such as face recognition [26], pedestrian detection [27–30], and crowd density estimation [31,32]. Camera et al. [33] summarized some pedestrian behavior analysis methods that can be used in artificial intelligence driving technologies and classified the methods according to whether pedestrian goals are considered and whether the pedestrian interactions are calculated. Based on the classification of the human trajectory prediction approaches, Rudenko et al. [34] showed that the neural network-based approaches can contribute to the discovery of statistical patterns inherent in pedestrian motion. Neural network methods can be highly efficient and can accurately predict pedestrian-related parameters; however, the successful implementation of such methods requires a comprehensive foundation, including an appropriate network architecture, large number of labelled samples, and sufficient training time. This article focuses on the review of pedestrian dynamics research and does not discuss the pedestrian image recognition techniques.

The use of neural networks in pedestrian dynamics research is an integration of computer science with traffic and safety domains. Using this advantageous new algorithm is important for research in the field of pedestrian dynamics. First, it can help us to understand human behavior in urban environments. This understanding can then be used to improve transportation, emergency management, security, and urban planning. For example, by understanding how pedestrians react to different stimuli, better crowd management techniques can be designed and more effective evacuation plans can be developed. Second, further pedestrian dynamics prediction work can be used to prevent accidents and disasters. For example, we can use neural network method to predict the crowd flow and identify potential bottlenecks in pedestrian traffic, even more to simulate how crowds will react to emergencies, such as fires or explosions.

Predicting pedestrian dynamics is a complex project because different scenarios have different characteristics that necessitate the use of different methods. It is important to understand the mechanism of employing neural networks and the types of pedestrian scenarios that can be addressed, from the previous studies. Firstly, it is important to understand the network configuration established in the previous studies. Since not every neural network method can effectively handle the pedestrian problem, the network architecture, activate functions, and datasets are crucial. An inappropriate neural network setup can lead to non-convergence of training data or poor accuracy of the prediction results, despite having provided adequate data for training. Secondly, the neural networks can generate only the potential associations for the given inputs and outputs. To summarize the previous considerations, this will be useful for the selection of appropriate input factors for the investigation of a new problem. Thirdly, benchmarks are extremely important for neural network research. Comparison between the models trained with the same datasets is an essential indicator for model evaluation. Therefore, researchers first need a comprehensive review, to be able to generate fast and accurate neural network models for the specific pedestrian problems.

This review is organized as follows. In the following sections, the authors sequentially present the studies on pedestrian trajectory prediction and pedestrian behavior prediction. For each category, the authors provide a detailed comparison and discussion, both conceptually and from the perspectives of methodology, datasets, and results. Finally, a short summary presents the development and outlook of the neural network methods in the field of pedestrian dynamics research.

2. Pedestrian trajectory prediction

2.1. Research topics

Pedestrian trajectory prediction has been a research hotspot and that has attracted considerable attention in the field of computer vision in the recent years. The application of pedestrian trajectory prediction is closely related to new technologies, such as autonomous driving and crowd recognition, and it contributes to the development of smart cities. With the popularity of urban video surveillance and the advancement of deep learning technology, an increasing number of new data-driven models have been proposed to predict the pedestrian trajectories and behaviors. These pedestrian trajectory prediction methods are more informative and intelligent than those in the traditional pedestrian dynamics studies, and they will probably become a new trend in the development of pedestrian dynamics and evacuation research.

The development of data-driven models requires large amounts of tagged data. By processing the pedestrian walking videos and tagging the coordination of each point on the movement path, the datasets are obtained and used as the model input. To compare the performance of different models, researchers have trained data-driven models with the same dataset sample and adopted a consistent measurement to establish a baseline for results analysis.

2.2. Methodology

Traditional models (such as social force model) can perform the trajectory prediction without labelled data; whereas, they are generally less accurate than the data-driven approaches. Due to the time-sequential nature of the video data and the fact that pedestrian trajectory is highly correlated with the previous movement path, RNN-based models can be one of the best choices to tackle this problem. Based on the previous efforts and tests, it is found that the LSTM-based models demonstrate an adequate performance on this issue.

Fig. 1 summarizes the classification of the existing pedestrian trajectory prediction models in terms of whether the impact of other pedestrians and the influence of scenes are considered. And Fig. 2 shows the classification of models in chronological order. The linear model is the simplest model, and the LSTM is the most basic data-driven model. Pedestrian interactions and physical influence are not considered in these two models, although it has a significant impact. Subsequent studies have focused on the influence of surrounding pedestrians on the motion trajectory. The most representative ones are the reapplication of the social force model and the proposal of the Social-LSTM model. Various other models are also proposed based on the concept of social-pooling from the Social-LSTM model. For example, the Group-LSTM improves the operation in social pooling to implement the trajectory prediction of different crowd groups. Recent studies pay special attention to the physical influences in the crowd scene or the interaction of distant pedestrians. The LSTM-based models are becoming more diverse, functionally refined, and are capable of handling highly complex pedestrian scenarios.

In the following text, the authors will describe the trajectory prediction methods in detail, particularly, the LSTM and Social-LSTM model, based on which a number of new models have been developed.

Modeled method:

•IGP (Interacting Gaussian Processes) [35]

This is a nonparametric statistical model based on the dependent output Gaussian processes, coupled through a nonlinear interaction. The information about the destination of a pedestrian is used in this model.

•SF (Social Force) [4,22]

Social force model is proposed by Helbing in 1995, the factors of grouping, collision, and attraction of individual pedestrians are considered, as well as social group affinity and predicted destinations have been modeled. As a classical pedestrian dynamics model, it has also been used as a trajectory prediction technique in some studies.

•Linear [22,24]

In this model, the pedestrian is assumed to move in a straight line; therefore, a straight line with the minimum mean squared error is calculated based on a regression of the pedestrian's trajectory, where the Kalman filter is the commonly used method to track and predict the next state of the pedestrian.

•LTA (Collision avoidance)

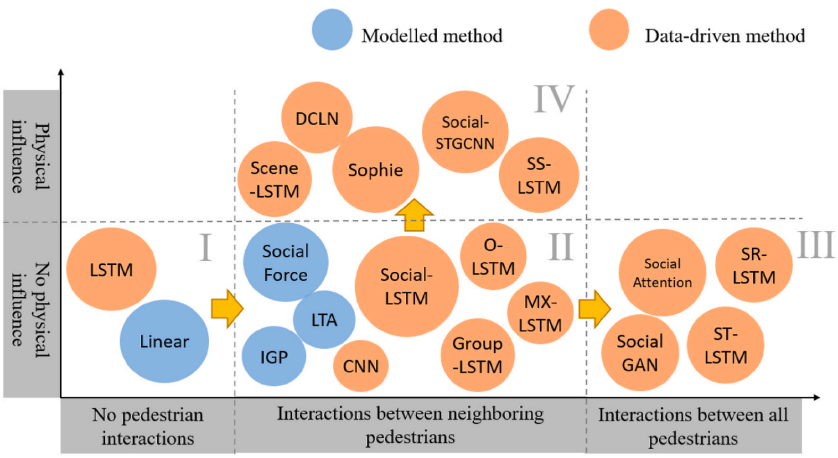

Fig. 1. Classification diagram of models on the topic of trajectory prediction.

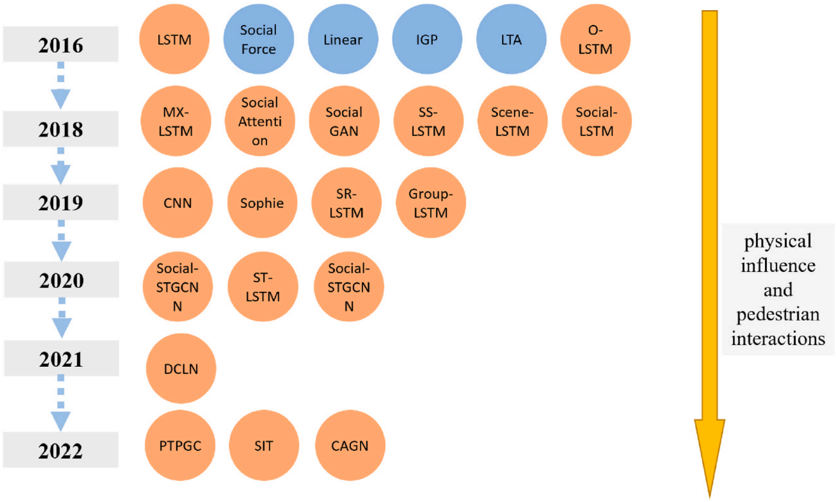

Fig. 2. Classification of models and time of first appearance.

This is a simplified version of the social force model that utilizes only the collision avoidance energy and is commonly referred to as linear trajectory avoidance.

Data-driven method:

•LSTM (Long-Short Term Memory) [22,36]

LSTM is a specialized RNN that is designed to solve the gradient vanishing and gradient explosion problems posed by long sequences. Compared with RNN, LSTM has two state vectors: cell state and hidden state. The cell state is a cumulative variable that changes gradually, while the hidden state is closely related to the current input and it can change at a faster rate. The architecture of the LSTM is shown in Fig. 3.

The operation of a common LSTM unit is controlled by three gates, that is, a forget gate z^f, an input gate z^i, and an output gate z^o. They can be calculated according to (1), where the symbol "+" represents the operation of concatenating two vectors, φ represents the activation function, W represents the weight, and b represents the bias.

$$\begin{cases} z^i = \varphi\left(W^i\left(x^t + h^{t-1}\right) + b^i\right) \\ z^i = \varphi\left(W^i\left(x^t + h^{t-1}\right) + b^i\right) \\ z^f = \varphi\left(W^f\left(x^t + h^{t-1}\right) + b^f\right) \end{cases} \tag{1}$$

Taking advantage of the conditioning effect of the gates, the LSTM computes the state variables and is able to output reasonable predictions, which are calculated according to (2), where, the symbols "$\oplus$" and "$\odot$" represent the operation of addition and multiplication of the corresponding elements of the vector, respectively.

$$\begin{cases} z = \tanh\left(W\left(x^t + h^{t-1}\right) + b\right) \\ c^t = z^f \odot c^{t-1} \oplus z^i \odot z \\ h^t = z^o \odot \tanh(c^t) \\ y^t = \varphi(W' h^t + b') \end{cases} \tag{2}$$

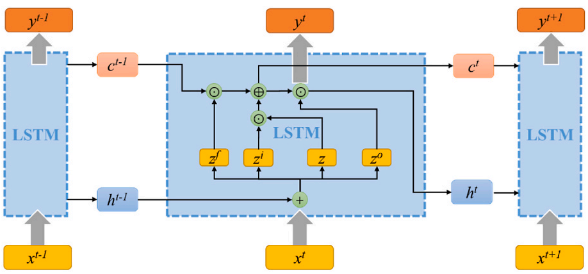

Fig. 3. Illustration of LSTM architecture; x^t, y^t, c^t, and h^t represent the input, output, cell state, and hidden state at time step t, respectively.

The problem of pedestrian trajectory prediction can be particularly regarded as training a long sequence. When the LSTM method is applied to trajectory prediction, it is assumed that each pedestrian is an independent LSTM. The coordinate points on each pedestrian trajectory comprise the input sequence x^t, and the corresponding prediction trajectory is the output sequence y^t. The social factors between the pedestrians are ignored in native LSTM; therefore, the pedestrian movement trajectories do not affect each other in the hidden layer. The predicted trajectory of a pedestrian is only related to the preceding movement path. The adequate performance of LSTM models in solving the problem of trajectory prediction provides a reliable foundation for Social-LSTM and its derivative methods.

•Social-LSTM [22]

Social-LSTM is an further development of the LSTM model. It overcomes the drawback of LSTM that cannot handle the interactions between pedestrians. Introducing the structure of social pool, where the hidden layer is shared, the LSTMs representing each pedestrian in a scene are connected in this architecture. Therefore, the influence of neighboring pedestrians on the movement trajectories can be modeled. An overview of Social-LSTM is shown in Fig. 4.

In the Social-LSTM method, the neighboring pedestrians are first placed into an $N_0 \times N_0$ grid at every time step according to their relative coordinates. A hidden-state tensor, H_t^i, is used to assemble the hidden state values of all the surrounding pedestrians in the grid, which is expressed as (3).

$$H_t^i(m,n,:) = \sum_{j \in N_i} l_{mn}\left[x_t^j - x_t^i, y_t^j - y_t^i\right] h_{t-1}^j \tag{3}$$

Integrating H_t^i into the original LSTM model, the Social-LSTM model can be constructed. In this model, the hidden state, h_t^i, plays an important role; further, it inherits the characteristics of the current pedestrian's movement path. Furthermore, it couples with the coordinate information, forming the tensor H_t^i, thereby affecting the trajectory prediction results of LTSM, which can be described as (4).

$$\begin{cases} e_t^i = \varphi\left(x_t^i, y_t^i; W_e\right) \\ a_t^i = \varphi\left(H_t^i; W_e\right) \\ h_t^i = \mathrm{LSTM}\left(h_{t-1}^i, e_t^i, a_t^i; W_l\right) \end{cases} \tag{4}$$

•LSTM (Occupancy map pooling LSTM) [22]

This is a reduced version of the Social-LSTM model that only pools the coordinates of the neighbors at every time instance and does not require any joint back-propagation across all trajectories during training. Therefore, this model would not be able to smoothly change paths to avoid future collisions.

- MX-LSTM (Mixing LSTM) [23]

MX-LSTM is an extension of Social-LSTM that takes into account the pedestrian head pose and visual field. By adding vislets (short sequences of head pose estimations), the view frustum of attention (VFOA) feature was integrated in this model. Pedestrians out of the VFOA were excluded from the social pool. By weakening the influence of the out-of-sight human on pedestrian movement, this model improved the pedestrian trajectory prediction performance.

- Social Attention [37]

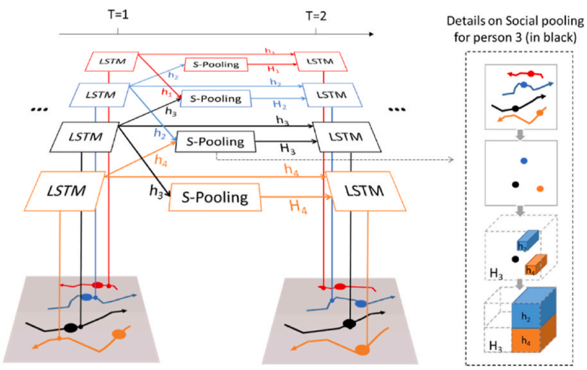

Fig. 4. Overview of the Social-LSTM method [22].

The social attention model is a trajectory prediction model based on the concept of the Social-LSTM method, which captures the relative importance of each person when navigating in the crowd, regardless of their proximity. Compared with the Social-LSTM method, the social attention model considered the interaction of pedestrian from a far distance, based on their velocities and advance. Therefore, this model can result in more accurate predictions.

•Social-GAN [38]

Social-GAN is a machine learning method that combines the Social-LSTM and generates adversarial networks. In comparison to the Social-LSTM approach, this considers the effect of those distant pedestrians in its hidden layer. Furthermore, a simple variety loss is introduced in this model to encourage diversity among the predicted samples.

•SS-LSTM (Social-Scene-LSTM) [25]

SS-LSTM considers the impact of scene on the pedestrian trajectory prediction. It uses three different LSTMs to capture the person, social, and scene scale information. The person-scale LSTM encoder observes the pedestrian's motion trajectory. The social-scale LSTM encoder processes the circular occupancy map, rather than the rectangular map in the Social-LSTM method. Further, the scene-scale LSTM encoder handles information based on the surrounding environments. Combining the output information of the three encoders and passing it through another layer of the LSTM encoder, the predicted pedestrian trajectory can be obtained.

•Scene-LSTM [24]

Scene-LSTM is another pedestrian trajectory prediction method that incorporates the scene information. By dividing the scene into multiple grid cells, the learning method is applied on each grid cell; therefore, it can predict the pedestrian trajectory based on the past movement of people in these grid cells [24]. proposed two types of Scene-LSTM method with the similar results, where, the Scene-LSTM-a uses the scene data from all the grid cells to predict the movement of people, and Scene-LSTM-n uses only the scene data from the non-linear grid cells.

•Social-Grid LSTM [39]

Social-Grid LSTM integrates the social pooling module and the two-dimensional Grid LSTM module. Furthermore, because of the use of the two-dimensional grid, the accuracy of the prediction is significantly improved, compared with the original Social-LSTM model.

•CNN [40]

The CNN method uses highly parallelizable convolutional layers to handle temporal dependencies. This model contains at least four hidden layers, and the convolutions can be easily parallelized. It can provide a better inference time performance and output the trajectory prediction results.

•Group-LSTM [41]

Group-LSTM method recognizes the similarity between the trajectories of groups pedestrians, whose interactions should not be processed in the social pool. By combining the coherent filtering algorithm with the LSTM networks, pedestrians are divided into several groups, and the impact of pedestrians in the same group is excluded from the social pool.

•SoPhie [42]

SoPhie is a GAN-based model that considers the pedestrian scene information. A visual attention model is used to process the static scene context, and an attentive model is adopted to cope with the movement trajectory of the pedestrians. Thereafter, a GAN module is used to learn the scene and trajectory data and output the probability distribution over each pedestrian's future path.

•SR-LSTM [43]

SR-LSTM stands for states refinement model for LSTM that emphasizes that the agent should be aware of the updated movement features and motion of the surrounding pedestrians. This model performs the interaction behavior between the pedestrians by creating a message passing framework. Through multiple processing refinement, it is ensured that each pedestrian agent can get the updated surrounding information.

•StarNet [44]

StarNet uses a star topology to represent the connections and influence of the pedestrians in motion. It consists of a unique centralized hub network and multiple host networks, and based on these, it can learn complicated interactions between the pedestrians and predict the future trajectories in an extremely short time.

•Social Affinity LSTM [45]

Social Affinity LSTM is proposed based on the idea of social-affinity map and LSTM method. By connecting the hidden states of the pedestrian and neighbors, a weight matrix corresponding to the bins of the social-affinity map, the trajectory and influence of each pedestrian can be predicted.

•Social-BiGAT [46]

Social-BiGAT is a graph-based generative adversarial network that generates trajectory predictions by modelling the social interactions of the pedestrians in a scene. In this model, all the pedestrians within the scene can interact and the pedestrian behaviors are represented by constructing a reversible mapping between the output trajectories and the latent. The experimental results confirm that this model outperforms the previous models.

•LSTM-based sequence-to-sequence trajectory prediction [47]

LSTM-based sequence-to-sequence trajectory prediction method adopts the LSTM encoder-decoder architecture, which allows the generation of a sequence with arbitrary length from a given sequence. And the method is aim to solve the problem of error accumulation over multiple time steps generated in traditional LSTM network.

•Social-STGCNN [48]

Social-STGCNN stands for the social spatio-temporal graph convolutional neural network model. It substitutes the need of aggregation methods by modeling the interactions as a spatio-temporal graph, whose edges model the social interactions between the pedestrians. This model manipulates the spatio-temporal graph using a graph CNN and temporal CNN, and it outperforms the previously discussed methods in terms of prediction error, computation time, and number of parameters.

•ST-LSTM [49]

ST-LSTM is another LSTM-based algorithm based on the method of Spatio-Temporal Graph, and it is aimed at tackling the pedestrian trajectory prediction of cluttered scene. In this model, the spatio-temporal convolutional block is introduced to extract features from pedestrians that combines two temporal convolutional networks and one spatial convolutional network. This model makes full use of the pedestrian spatial features and is able to output similar predictions to a real-life situation.

•DCLN [49]

The Deep Convolutional LSTM Network (DCLN) model is able to represent the complex, multi-channel spatial features surrounding pedestrians. In contrast to the traditional LSTM-based methods that use 1D vector to represent the environmental information in the LSTM network, the DCLN adopts a multi-channel tensor to represent the environmental information such that more accurate trajectory prediction results can be obtained.

•PTPGC [50]

“ PTPGC is a model based on graph attention and convolutional long short-term memory. The feature of this model is that it considers pedestrian interactions in both encoding stage and trajectory prediction stage. Moreover, the distributions of the trajectories generated by this model are more converged compared to other methods.

•SIT [51]

Social Interpretable Tree (SIT) employs a coarse-to-fine optimization strategy to build the path tree based on the observed pedestrian trajectory data. Different from other methods, SIT embed the future pedestrian trajectories into a discrete structure space, and it has a better performance on long-term prediction work.

•CAGN [52]

Complementary attention gated network (CAGN) focus more on the peculiar pedestrian scenarios, and this method proposes a dual-path architecture including normal and inverse attention. Benefiting from its complementary attention mechanism on spatial and temporal dimension, the model can handle diversified pedestrian trajectory prediction task in more complex scenarios.

2.3. Datasets

There are several publicly available pedestrian trajectory datasets: New York Grand Central (GC) [53], the ETH Dataset [54], UCY Dataset [55], CUHK Crowd Dataset [56], MIT Traffic Dataset [57], Town Center Dataset [58], and Stanford Drone Dataset (SDD) [59], shown as Fig. 5. These datasets contain a variety of crowd behaviors, such as bi-direction flows, path reversals, groups walking together, groups crossing each other, and pedestrian conversations. Among these, ETH and UCY are the two most widely used datasets.

•ETH [54]

The ETH dataset consist of two scenes (ETH and HOTEL). The ETH scene has a total of 1448 identified frames with 367 pedestrians as well as 61 crowd groups, and the HOTEL scene has a total of 1430 frames, with 420 pedestrians and 41 crowd groups recognized. The ETH dataset is available at https://icu.ee.ethz.ch/research/datsets.html.

•UCY [55]

The UCY dataset contains two scenes and three components (UNIV, ZARA01 and ZARA02). This dataset reflects the trajectory of the pedestrians in a campus or city street area, and the three video components recorded 434, 148, and 204 pedestrian tracks, respectively. The UCY dataset is available at https://graphics.cs.ucy.ac.cy/research/downloads/crowd-data.

•CUHK Crowd

CUHK Crowd dataset contains 474 videos from 215 crowded scenes, and it has been used in the papers [56,60]. The dataset is available at https://amandajshao.github.io/projects/CUHKcrowd_files/cuhk_crowd_dataset.htm.

•MIT Traffic Dataset [57]

The dataset includes a video and trajectories collected from the New York Grand Central Station, which is available at http://mmlab.ie.cuhk.edu.hk/project/dynamicagent/

•Oxford Town Center Dataset [58]

The Oxford Town Center dataset is a CCTV video of pedestrians in a downtown area. It has been used for face recognition system, social distancing algorithms and human sex classification. The dataset is apparently no longer being distributed.

•Stanford Drone Dataset [61]

The Stanford Drone Dataset is a large dataset, whose total size reaches about 69 GB. It contains various types of agents including pedestrians, bicyclists, skateboarders, cars, and buses. The dataset is available at https://cvgl.stanford.edu/projects/uav_data/

2.4. Measurements

Average Displacement Error (ADE) and Final Displacement Error (FDE) are the measurements widely used to compare the difference between the predicted and ground truth trajectories.

ADE is the mean square error, as introduced in Pellegirini et al. that indicates the error over all the estimated points of a predicted trajectory, and the ground truth path points over the period from $t_{obs} + 1$ to t_{pred}. The ADE is calculated as (5).

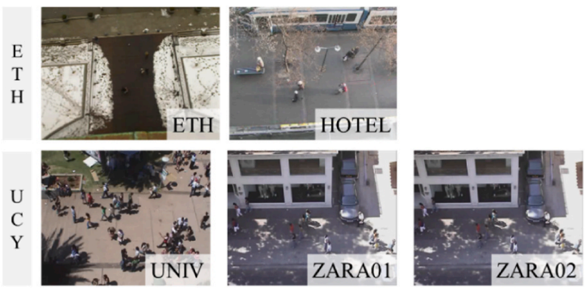

Fig. 5. Video frame captures from ETH and UCY datasets.

$$ADE = \frac{\sum_{i=1}^{n}\sum_{t=t_{\text{obs}}+1}^{t_{\text{pred}}}\left\|\boldsymbol{X}_{i,t}^{\text{pred}} - \boldsymbol{X}_{i,t}^{\text{obs}}\right\|}{n\left(t_{\text{pred}} - \left(t_{\text{obs}} + 1\right)\right)} \tag{5}$$

The FDE indicates the error of predicted destination and the true destination of the prediction, which can be calculated as (6). In other words, it is only concerned with the final discrepancy.

$$FDE = \frac{\sum_{i=1}^{n}\left\|\boldsymbol{X}_{i,T_{\text{pred}}}^{\text{pred}} - \boldsymbol{X}_{i,T_{\text{pred}}}^{\text{obs}}\right\|}{n} \tag{6}$$

Average non-linear displacement error is, in fact, the ADE at the non-linear region of a trajectory because multiple errors in the trajectory prediction occur in this region. The non-linear region can be identified by delineating the second derivative threshold. This method has also been used to measure the accuracy of the model in few studies.

Table 1
Quantitative results obtained based on the different methods for five dataset scenes.

Method	Reference	ADE (m)						FDE (m)					
		ETH	HOTEL	UNIV	ZARA1	ZARA2	AVG	ETH	HOTEL	UNIV	ZARA1	ZARA2	AVG
Linear	[22]	0.80	0.39	0.57	0.47	0.45	0.53	1.31	0.55	0.89	0.91	1.14	0.97
	[24]	0.17	0.21	0.23	0.36	0.15	0.22	0.24	0.25	0.03	0.32	0.09	0.19
	[46]	1.33	0.39	0.82	0.62	0.77	0.79	2.94	0.72	1.59	1.21	1.48	1.59
SF	[22]	0.41	0.25	0.48	0.40	0.40	0.39	0.59	0.37	0.78	0.60	0.68	0.60
	[23]	–	–	2.57	2.88	2.32	2.59	–	–	4.62	5.55	4.35	4.84
LTA	[22]	0.54	0.38	0.51	0.37	0.40	0.44	0.77	0.64	0.95	0.66	0.72	0.74
	[23]	–	–	2.49	2.74	2.23	2.49	–	–	4.66	5.55	4.35	4.85
IGP	[22]	0.20	0.24	0.61	0.39	0.41	0.37	0.43	0.37	1.82	0.39	0.42	0.69
LSTM	[22]	0.60	0.15	0.52	0.43	0.51	0.44	1.31	0.33	1.25	0.93	1.09	0.98
	[23]	–	–	0.67	0.90	1.09	0.89	–	–	1.39	1.85	2.15	1.80
	[37]	0.59	0.35	0.25	0.38	0.40	0.39	5.28	4.42	1.55	3.57	6.39	3.84
	[24]	0.16	0.21	0.27	0.33	0.19	0.23	0.29	0.23	0.05	0.28	0.21	0.21
	[39]	0.13	0.16	0.15	0.28	0.37	0.25	0.22	0.28	0.65	0.23	0.51	0.38
	[46]	1.09	0.86	0.61	0.41	0.52	0.70	2.31	1.91	1.31	0.88	1.11	1.52
	[44]	0.70	0.55	0.36	0.25	0.31	0.43	1.45	1.17	0.77	0.53	0.65	0.91
	[45]	0.04	**0.01**	0.08	0.06	0.02	0.04	0.05	0.03	0.13	0.04	0.05	0.06
Social-LSTM	[22]	0.50	0.11	0.27	0.22	0.25	0.27	1.07	0.23	0.77	0.48	0.50	0.61
	[23]	–	–	0.62	0.68	0.63	0.64	–	–	1.12	1.31	0.79	1.07
	[37]	0.46	0.42	0.21	0.41	0.36	0.37	4.55	3.57	0.65	3.39	4.45	3.32
	[24]	0.18	0.25	0.25	0.37	0.19	0.25	0.34	0.29	0.03	0.32	0.10	0.22
	[39]	0.06	0.11	0.34	0.20	0.24	0.19	0.13	0.24	0.34	0.37	0.55	0.33
	[44]	0.73	0.49	0.41	0.27	0.33	0.45	1.48	1.01	0.84	0.56	0.70	0.91
	[45]	0.03	**0.01**	**0.01**	**0.01**	**0.01**	0.02	0.03	**0.02**	0.06	**0.02**	0.02	**0.03**
	[49]	1.09	0.79	0.67	0.47	0.56	0.72	2.35	1.76	1.40	1.00	1.17	1.54
O-LSTM	[22]	0.49	0.09	0.35	0.22	0.28	0.28	1.06	0.20	0.90	0.46	0.58	0.64
MX-LSTM	[23]	–	–	0.49	0.59	0.35	0.48	–	–	1.12	1.31	0.79	2.30
Social Attention	[37]	0.39	0.29	0.20	0.30	0.33	0.30	3.74	2.64	0.52	2.13	3.92	2.59
	[24]	0.44	0.44	0.19	0.47	0.56	0.42	0.50	0.55	0.10	0.62	0.71	0.50
	[44]	2.45	2.19	2.92	1.66	2.30	2.30	5.78	4.94	4.95	2.64	4.75	4.81
Social-GAN	[38]	0.81	0.72	0.60	0.34	0.42	0.58	1.52	1.61	1.26	0.69	0.84	1.18
	[44]	0.61	0.48	0.36	0.21	0.27	0.39	1.22	0.95	0.75	0.42	0.54	0.78
	[62]	0.75	0.48	0.36	0.21	0.27	0.39	1.22	0.95	0.75	0.42	0.54	0.78
SS-LSTM	[25]	0.10	0.07	0.08	0.05	0.05	0.07	0.24	0.12	0.13	0.08	0.08	0.09
Scene-LSTM-a	[24]	0.11	0.06	0.11	0.07	0.06	0.08	0.19	0.06	**0.02**	0.08	0.03	0.08
Scene-LSTM-n	[24]	0.10	0.06	0.09	0.07	0.05	0.07	0.18	0.07	**0.02**	0.07	**0.02**	0.07
Social-GridLSTM	[39]	0.06	0.06	0.27	0.16	0.19	0.15	**0.11**	0.09	0.52	0.29	0.33	0.27
CNN	[40]	0.59	1.04	0.57	0.43	0.34	0.59	1.17	2.07	1.21	0.90	0.75	1.22
Group-LSTM	[41]	0.28	0.28	0.56	0.23	0.34	0.34	1.12	0.89	1.48	0.91	1.49	1.18
SoPhie	[42]	0.70	0.76	0.54	0.30	0.38	0.54	1.43	1.67	1.24	0.63	0.78	1.15
SR-LSTM	[43]	0.63	0.37	0.51	0.41	0.32	0.45	1.25	0.74	1.10	0.90	0.70	0.94
Social-BiGAT	[46]	0.69	0.49	0.55	0.30	0.36	0.48	1.29	1.01	1.32	0.62	0.75	1.00
StarNet	[44]	0.31	0.46	0.21	0.25	0.26	0.30	0.54	0.91	0.40	0.47	0.53	0.57
Social-Affinity LSTM	[45]	**0.02**	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**	**0.02**	**0.02**	**0.05**	**0.02**	**0.01**	**0.03**
Social-STGCNN	[48]	0.64	0.49	0.44	0.34	0.30	0.44	1.11	0.85	0.79	0.53	0.48	0.75
ST-LSTM	[49]	0.57	0.41	0.38	0.31	0.33	0.40	1.02	0.73	0.84	0.47	0.51	0.72
DCLN	[62]	0.07	0.03	0.06	0.02	0.02	0.05	0.12	0.04	0.08	0.04	0.03	0.06
PTPGC	[50]	0.62	0.25	0.40	0.29	0.27	0.37	1.01	0.33	0.70	0.49	0.47	0.60
SIT	[51]	0.39	0.14	0.27	0.19	0.16	0.23	0.62	0.22	0.47	0.33	0.29	0.38
CAGN	[52]	0.41	0.13	0.32	0.21	0.16	0.25	0.65	0.23	0.54	0.38	0.33	0.43

2.5. Discussion

Alahi et al. [22] were the first to investigate the LSTM-based trajectory prediction model, and their study paved the way for other researchers. Since the Social-LSTM model has become the baseline that is generally compared with the other models, almost all of the studies have set a similar sample configuration based on it. Table 1 summarizes the experimental results on the two datasets (five scenes) in the existing literature. These studies predicted trajectory for 12 frames (4.8 s) using the observed trajectories of 8 frames (3.2 s).

As shown in the table, the results may vary across studies, despite using the same approach and the same dataset. It is well-known that the training of neural network models is inherently random and produces diverse results. In addition, the authors believe there are two other reasons for this variation. First, although the datasets are identical, different researchers may have selected different data samples for model training, which could lead to some discrepancies in the experimental results. Second, few scholars have published their experimental codes and dataset, while others have not. In some cases, the researchers could replicate the model only according to the description in the previous literature that led to the differences.

Due to the large variation of results provided by the different studies, it is difficult to make an explicit quantitative comparison. However, it can still be observed that the models, in general, consider that a particular scene has more influence over the accuracy in the trajectory prediction. Table 1 shows the methods used in different literature and their computational results. The model with the best performance for each dataset and average error is emboldened and underlined. Such a simple comparison is only informative, considering the differences between studies. It indicates that there is still a need to establish a comprehensive and standardized evaluation methodology to provide an accurate assessment of these prediction models.

Fig. 6 shows the distribution of ADE and FDE results in the previous studies. In most cases, the final prediction error in the next 3.2 s is no more than 2 m, which may be beyond the acceptable range. Among them, the DCLN, Scene-LSTM, and SR-LSTM models exhibit an outstanding performance and the error can be controlled within 0.2 m. Except for a few case studies [37], it is also observed that the ratio of FDE and ADE is approximately 2.0 that is exactly the slope of the scatter plot. Moreover, the FDE is twice as large as ADE, indicating that the trajectory prediction errors of these methods can be linearly additive over time. From the figure, it can be observed that the methods in region III of Fig. 6 have better prediction accuracy, which indicates that calculating the interactions between all pedestrians can significantly improve the prediction.

In addition, few studies on trajectory scene did not use the ETH and UCY datasets for training, owing to the difference on the research content; therefore, the effect of these models cannot be effectively compared. Yi et al. (2016) proposed a model, called Behavior-CNN. The factors of location map, location awareness property, semantic meanings of learned filters, and influence of receptive fields are incorporated in this model; therefore, the walking behavior of pedestrian in various environments can be learned and trained efficiently [63]. proposed a path prediction model using the object attributes and semantic environment. The model considers a variety of area types in the scene, including the areas such as, roads, trees, buildings, and travel lanes. In addition, the pedestrians are classified into normal and cycling pedestrians. Such processing enables the application of LSTM-based methods to complicated street scenes [64]. proposed a pedestrian trajectory prediction method called location-velocity attention model. In this model, the location and velocity information are considered as input data, and therefore, there is no need to provide additional context of surrounding pedestrians and scene information. Moreover [65], proposed a new graph neural network method that combines the recurrent neural network [66] and graph neural network [67]. The study indicates that the rationale behind the interactions between pedestrians is important in trajectory prediction. Although these models are difficult to be compared cross-sectionally with those of the existing studies, they provide unique approaches to the pedestrian-trajectory prediction in specific scenarios.

Overall, current neural network models for pedestrian trajectory prediction can mainly achieve short-term trajectory prediction, roughly the pedestrian's path in the next 10 s or less. And long-term crowd trajectory prediction is necessary for the safety management of crowd flow, and it can also contribute to the safety evacuation prediction. In the future, the development of pedestrian trajectory prediction models can be considered from several perspectives: First, consider that how to identify the characteristics of the surrounding input information, so that the prediction results have a higher accuracy. Second, if the necessary and streamlined neural network structure is determined, so that the model has a good response speed. Third, the current model is limited in applications by the

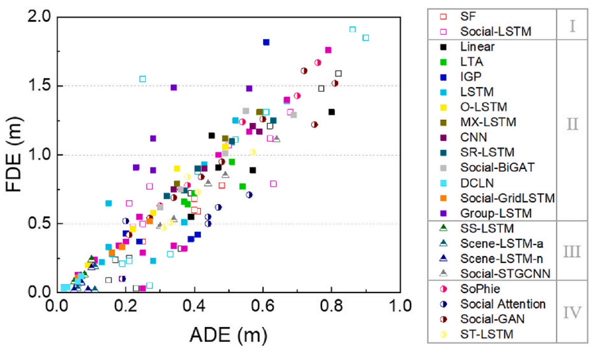

Fig. 6. Distribution of ADE and FDE of the different prediction models.

demand for computational resources, and the future model must be able to realize a longer spatial and temporal range for trajectory prediction, which can help pedestrian safety management in real life.

3. Pedestrian behavior prediction

3.1. Research topics

There are many research topics related to human psychology and pedestrian behavior, such as target selection and walking-direction decision. In the previous studies, the pedestrian behavior prediction has generally been presented as a decision-making process according to the information on the individual needs and surrounding conditions of the pedestrians. The main factors affecting pedestrians' decision making, including walking status, environmental conditions, and destination, are illustrated in Fig. 7.

The prediction of pedestrian behavior was more focused on human decisions at a certain moment, while the historical information of pedestrians was given less importance. Unlike the pedestrian trajectory studies, behavior predictions were usually not time series-based, and the pedestrian status, the surrounding environmental conditions as well as the stimulus that affects pedestrians have been focused on. Due to the fact that time-series prediction is unnecessary, the deep neural network method would be an appropriate tool for solving problems. The deep neural network, also called the multilayer perceptron model (MLP), has been applied and performs satisfactorily in many pedestrian dynamics studies [68,69]. In the following sections, the MLP methods and the applications in pedestrian behavior prediction work will be introduced.

3.2. Methodology

The MLP is one of the most basic neural network models and its architecture is shown in Fig. 8. It is a type of fully connected neural network that contains an input layer, hidden layers, and an output layer.

Each layer contains a number of neurons. The number of neurons of the input layer and output layer are directly related to the input parameters and expected output parameters for the particular investigation scenario. The number of neurons in the hidden layers, N_h, are usually considered to be between the value of N_{in} and N_{out}. In few studies, the N_h is set according to (7),

$$N_h = \frac{N_{in} + N_{out}}{2} + \sqrt{N_s}, \tag{7}$$

where, N_{in} and N_{out} represent the number of neurons in the input and output layer respectively, and N_s represents the size of the training samples. Neurons of a layer are connected to the neurons of the next layer, and each connection carries weight information. The neurons of the former layer are treated as inputs, and the output is provided to the neurons of the next layer through activation functions. The relationship between the inputs x, outputs y, and activation functions φ can be described by (8).

$$y^i = \varphi\left(\sum\nolimits_j W^j x^j + b^i\right) \tag{8}$$

By training the MLP using a large amount of sample data, the parameters of the weights (W) and bias (b) are obtained based on the back propagation method. Using the trained neural networks, it is possible to conduct an outcome prediction of the out pedestrian dynamics problems.

Because the MLP model is simpler than the LSTM model, it is applicable to a wider range of pedestrian dynamics issues. However, the MLP model is not appropriate for time-series predictions. The MLP model is not structurally based on the RNN inside, and hence cannot store the previous prediction results. Therefore, MLP models are only capable of handling static problems in pedestrian dynamics.

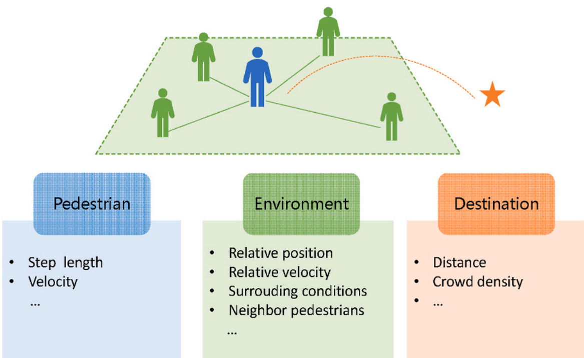

Fig. 7. Factors affecting pedestrian behaviors in the previous studies.

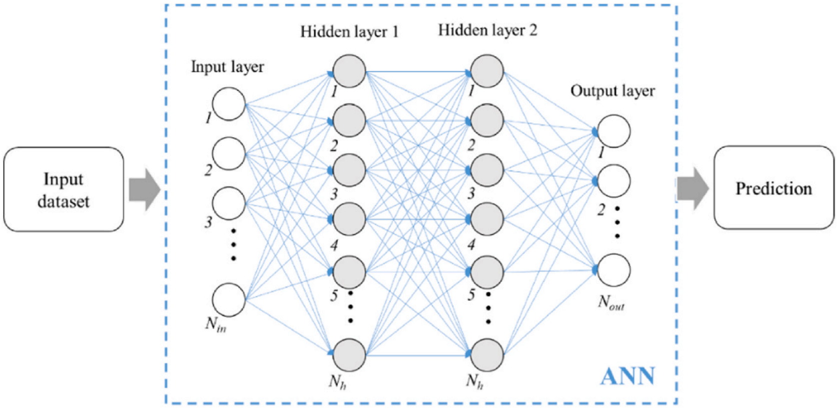

Fig. 8. Typical architecture of the multilayer perceptron (MLP) model.

3.3. Discussions

MLP has been applied to solve a variety of pedestrian behavior prediction scenarios, such as prediction of pedestrian velocity, exit choice of an area, and pedestrian step-length. Unlike trajectory prediction, the subjects of these studies are not the same, and therefore, the inputs, outputs, measurements, and datasets for these problems are different altogether. Table 2 summarizes few studies on pedestrian dynamics problems that have employed the MLP method that are comparative in methodology. To make a comprehensive comparison between these studies, the research topic, neural network configuration, dataset, and results are described. Specifically, the neural network configurations of previous studies are highlighted. The parameters that make up the input and output are listed, and the numbers within brackets indicate the use of multiple neurons. Based on the previous studies, the researchers have verified that the neural network configurations can produce satisfactory results corresponding to the respective research topics, and it is highly important to understand the methodological use of neural networks for the investigation of the other pedestrian problems.

Pedestrian movement simulation, particularly, the walking direction and velocity preferences of the pedestrians in a particular environment have been investigated by many scholars. Although deep neural networks can hardly achieve the appropriate results in the time-series predictions, they have a better extendibility for different pedestrian scenarios and research requirements. Shao et al. [70] tried to combine CA model with neural networks. By mapping the parameters of the neighborhood grids as input neurons and mapping the pedestrian motion direction to several output neurons, a neural network was built and verified to feature an adequate accuracy. This work discusses the method of transforming a grid model into a neural network. Ma et al. [69] investigated the pedestrian movement on the urban streets. It is noted that the movement of pedestrians affect the other pedestrians; therefore, using the current motion parameters and information of the nearest neighbor pedestrians, the next movement is appropriately predicted. Song et al. [71] trained the neural network with the experimental data of four pedestrian scenarios. The proposed model can generate more realistic results than the social force model; however, it still does not converge with the real-life data; particularly for the competitive situations. Zhao et al. [72] proposed a composite neural network to study the unidirectional and bidirectional pedestrian flow. They used two sets of neural networks to predict the magnitude and direction of the pedestrian speed separately. This is an effective way to cope with the complex problems that may involve a large neural network and lead to inefficient training. Moreover, Goldhammer et al. [68] conducted a pedestrian trajectory prediction research using polynomial least squares approximation, in conjugation with the MLP method. The lateral and longitudinal movements of pedestrians was treated separately in this model, and adequate prediction results were obtained for the sample data of pedestrian waiting, starting, walking, and stopping. In addition to the pedestrian speed, the problems of path finding as well as risk assessment can also be performed by the neural network method. Sharma et al. [73] investigated goal finding by intelligent agents based on a neural network approach. Yuen et al. [74] studied the influence of environmental factors on human decision making in urban subway scenarios. The exit preferences of pedestrians were predicted based on the data of crowd density distribution and exit distance. Zhao et al. [21] proposed a feature extraction and mapping network model for crowd evacuation scenarios. By using the fuzzy neural network method, the level of crowd stampede in a specific scenario can finally be output.

In general, the MLP method is more often applied to the macroscopic pedestrian research, particularly in prediction decision-making-related problems. Benefitting from the simple architecture, this method is user-friendly and has been widely used in many pedestrian scenarios. It can be used to prejudge human behaviors and quantify the risk level in some emergency evacuation situations. Many previous studies have shown that well-trained MLP models can produce satisfactory accuracy for the predicting results. However, there are some shortcomings to using MLP for pedestrian behavior prediction. First, the scale of the datasets used in previous studies is small, and only a minority of these datasets are publicly available. This makes it difficult to verify whether a neural network has been trained sufficiently. Second, the experimental results of previous studies are difficult to reproduce by other researchers. This is because the datasets used in these studies are often not available, and the experimental settings are not always well-documented. Additionally, there is no consensus yet on the type of pedestrian dynamics problems that can be investigated using MLP. This is because MLP is a general-purpose algorithm, and its performance can vary depending on the specific problem being solved.

Table 2
Studies on pedestrian walking behaviors based on MLP.

Shao et al. [70]	Research topic	Pedestrian simulation using CA model embedded by BP neural network
	Input layer (9)	• Current relative position (1) • Neighbor grads (4) • Surrounding conditions (4)
	Hidden layer	1 hidden layer
	Output layer (3)	Direction choice of a particular pedestrian (3)
	Dataset size	78 samples
	Results	Accuracy is 94.8 % out of 78 samples and 91.3 % out of 23 new samples
Yuen et al. [74]	Research topic	Passenger route choice prediction in a transportation station
	Input layer (9)	• shortest distance between entrance and the three gate groups (3) • number of gates in each gate group (3) • percentage of wide gate in each gate group (3)
	Hidden layer (27)	1 hidden layer (27)
	Output layer (3)	probability of passengers choosing gate group (3)
	Dataset size	617 samples (70 % training, 30 % testing)
	Results	Accuracy of their ensemble approach reached 86.17 %.
Yuen et al. [17]	Research topic	Passengers making a route choice between two escalators in a transportation station
	Input layer (6)	• walking velocity of passengers (1) • initial position of passengers (1) • initial density of two escalators (2) • final density of two escalators (2)
	Hidden layer (21)	1 hidden layer (21)
	Output layer (1)	passengers choosing escalator 1 or 2 (1)
	Dataset size	621 samples (50 % training, 25 % validation, 25 % testing)
	Results	Prediction accuracy: 86 %
Ma et al. [69]	Research topic	Pedestrian movement simulation
	Input layer (25)	• relative distance to the nearest pedestrians (10) • relative velocity to the nearest pedestrians (10) • current position of subject pedestrian (3) • current velocity of subject pedestrian (2)
	Hidden layer (81)	1 hidden layer (81)
	Output layer (2)	pedestrian velocity of next time step (2)
	Dataset size	4496 samples (80 % training, 15 % validation, 5 % testing)
	Results	The mean relative distance error was calculated to be 0.322. The lane-formation phenomenon can be reproduced by this approach.
Song et al. [71]	Research topic	Pedestrian movement simulation
	Input layer (22)	• pedestrian walking speed magnitude (1) • pedestrian velocity angle (1) • relative distance to the nearest 5 neighbors (10) • relative velocity to the nearest 5 neighbors (10)
	Hidden layer (110)	2 hidden layers (both 110)
	Output layer (2)	velocity of subject person at the next time step (2)
	Dataset size	61305 samples (90 % training, 10 % testing)
	Results	The ANN approach showed considerably lower position fluctuations and provided an improved speed-density curve than the social force model.
Ma et al.[75]	Research topic	Pedestrian walking behavior simulation (velocity and step length)
	Input layer (34)	• current position: distance from destination, left boundary, and right boundary (3) • x-component and y-component of current velocity (2) • step length of current footstep (1) • relative position of nearest neighbor pedestrian (14) • relative velocity of nearest neighbor pedestrian (14)
	Hidden layer (80)	1 hidden layer (80)
	Output layer (3)	• x-component and y-component of velocity in next footstep (2) • step length of next footstep (1)
	Dataset size	3813 samples (80 % training, 15 % validation, 5 % testing)
	Results	Mean of the learning performance reached 0.900, and it is considered that the DNN approach can appropriately predict the details of the local microscopic pedestrian walking behavior.
Zhao et al. [72]	Research topic	Pedestrian walking behaviors in both unidirectional and bidirectional flow
	Input layer	• relative distance between subject person and other pedestrians • pedestrian distribution of forward space (98)
	Hidden layer	• 2 hidden layers for magnitude prediction • 2 hidden layers for direction prediction (50 + 25)
	Output layer	• velocity magnitude of the subject pedestrian • velocity direction angle of the subject pedestrian (1)
	Dataset size	(80 % training, 15 % validation, 5 % testing)

(continued on next page)

Table 2 (*continued*)

Results	The mean distance error, mean trajectory error, and relative distance error can be controlled within 0.20 m in either unidirectional flow scenario or bidirectional flow scenario.

Despite these shortcomings, MLP is an optimal method for behavior prediction and evacuation decision-making prejudge. However, more research is needed to systematize the study of MLP methods and make the prediction of pedestrian behavior by employing neural networks more prevalent. This includes developing larger and more publicly available datasets, standardizing experimental settings, and conducting more comparative studies of different MLP methods.

4. Conclusions

In this article, studies on the pedestrian and evacuation dynamics based on neural network methods are reviewed. With the exception of visual image processing, the main research topics discussed are pedestrian trajectory prediction and pedestrian behavior prediction. Neural networks can be trained on data from real-world pedestrian movements and be used to analyze pedestrian flow in real time. The pedestrian recording data can be used to train the neural network to predict how pedestrians will behave in different situations. Pedestrian flow predictions and simulations can be further used to improve crowd management techniques, evacuation plans, and urban planning.

Pedestrian trajectory prediction is closely associated with computer vision, and many scholars from the field of computer science are engaged in this research. Among them, the LSTM-based method is the most widely used method that provides satisfactory predictions. However, current work can only produce accurate results within a few seconds. Such a time span does not yet meet the requirements for pedestrian evacuation analysis. Most of the studies on pedestrian behavior and evacuation prediction have been carried out by research groups who have been involved in pedestrian dynamics and public safety science for years. The main drawbacks of these studies are the lack of sufficiently large datasets and the lack if inter-comparison of results between the other researches. There is a need to focus on few common concerns and publish more datasets, and to develop comparable approaches accordingly. As the computing power increases, with more mature techniques based on neural networks and the increasing involvement of scholars, the outcome of researching these topics will be accelerated, finally contributing to the widespread applications of urban crowd safety management and smart city development.

CRediT authorship contribution statement

Shenshi Huang: Conceptualization, Formal analysis, Funding acquisition, Investigation, Writing – original draft. **Ruichao Wei:** Conceptualization, Funding acquisition, Investigation, Writing – original draft. **Liping Lian:** Writing – review & editing. **Siuming Lo:** Writing – original draft. **Shouxiang Lu:** Writing – original draft.

Declaration of competing interest

The authors declared that they do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgements

This work was supported by Guangdong Basic and Applied Basic Research Foundation (2023A1515110492), Scientific Research Startup Fund for Shenzhen High-Caliber Personnel of SZPT (No. 6022310043K and 6023330004K) and Postdoctoral Later-stage Foundation Project of Shenzhen Polytechnic (6022271007K).

References

[1] N. Shiwakoti, X. Shi, Z. Ye, A review on the performance of an obstacle near an exit on pedestrian crowd evacuation, Saf. Sci. 113 (November 2018) (2019) 54–67, https://doi.org/10.1016/j.ssci.2018.11.016.
[2] R. Lovreglio, M. Kinateder, Augmented reality for pedestrian evacuation research: promises and limitations, Saf. Sci. 128 (April) (2020) 104750, https://doi.org/10.1016/j.ssci.2020.104750.
[3] L. Fu, Y. Liu, P. Yang, Y. Shi, Y. Zhao, J. Fang, Dynamic analysis of stepping behavior of pedestrian social groups on stairs, J. Stat. Mech. Theor. Exp. 2020 (6) (2020) 063403.
[4] D. Helbing, P. Molnár, Social force model for pedestrian dynamics, Phys. Rev. 51 (5) (1995) 4282–4286, https://doi.org/10.1103/PhysRevE.51.4282.
[5] R. Mehran, A. Oyama, M. Shah, Abnormal crowd behavior detection using social force model, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, 2009 IEEE, 2009, pp. 935–942, https://doi.org/10.1109/CVPRW.2009.5206641, 1.
[6] F. Weifeng, Y. Lizhong, F. Weicheng, Simulation of bi-direction pedestrian movement using a cellular automata model, Phys. Stat. Mech. Appl. 321 (3–4) (2003) 633–640, https://doi.org/10.1016/S0378-4371(02)01732-6.
[7] D. Li, B. Han, Behavioral effect on pedestrian evacuation simulation using cellular automata, Saf. Sci. 80 (2015) 41–55, https://doi.org/10.1016/j.ssci.2015.07.003.
[8] Y. Li, et al., A review of cellular automata models for crowd evacuation, Phys. Stat. Mech. Appl. 526 (November) (2019) 120752, https://doi.org/10.1016/j.physa.2019.03.117.
[9] D. Xu, et al., Simulating multi-exit evacuation using deep reinforcement learning, Trans. GIS 25 (3) (2021) 1542–1564, https://doi.org/10.1111/tgis.12738.

[10] H.W. Fischer, et al., Evacuation behaviour: why do some evacuate, while others do not? A case study of the Ephrata, Pennsylvania (USA) evacuation, Disaster Prev. Manag. 4 (4) (1995) 30–36, https://doi.org/10.1108/09653569510093414.
[11] W. mei Gai, Y. Du, Y. feng Deng, Evacuation risk assessment of regional evacuation for major accidents and its application in emergency planning: a case study, Saf. Sci. 106 (March) (2018) 203–218, https://doi.org/10.1016/j.ssci.2018.03.021.
[12] N. Pelechano, A. Malkawi, Evacuation simulation models: challenges in modeling high rise building evacuation with cellular automata approaches, Autom. ConStruct. 17 (4) (2008) 377–385, https://doi.org/10.1016/j.autcon.2007.06.005.
[13] R.Y. Guo, H.J. Huang, S.C. Wong, A potential field approach to the modeling of route choice in pedestrian evacuation, J. Stat. Mech. Theor. Exp. 2013 (2) (2013), https://doi.org/10.1088/1742-5468/2013/02/P02010.
[14] Y. Hu, Research Review of China Emergency Evacuation, 23, Icpel, 2017, pp. 185–188, https://doi.org/10.2991/icpel-17.2017.48.
[15] M. Liu, S.M. Lo, The quantitative investigation on people's pre-evacuation behavior under fire, Autom. ConStruct. 20 (5) (2011) 620–628, https://doi.org/10.1016/j.autcon.2010.12.004.
[16] S.M. Lo, et al., An artificial neural-network based predictive model for pre-evacuation human response in domestic building fire, Fire Technol. 45 (4) (2009) 431–449, https://doi.org/10.1007/s10694-008-0064-6.
[17] J.K.K. Yuen, E.W.M. Lee, W.W.H. Lam, An intelligence-based route choice model for pedestrian flow in a transportation station, Applied Soft Computing Journal 24 (2014) 31–39, https://doi.org/10.1016/j.asoc.2014.05.031.
[18] U. Atila, et al., SmartEscape: a mobile smart individual fire evacuation system based on 3D spatial model, ISPRS Int. J. Geo-Inf. 7 (6) (2018), https://doi.org/10.3390/ijgi7060223.
[19] K. Tkachuk, X. Song, I. Maltseva, Application of artificial neural networks for agent-based simulation of emergency evacuation from buildings for various purpose, in: IOP Conference Series: Materials Science and Engineering, 365, 2018, https://doi.org/10.1088/1757-899X/365/4/042064, 4.
[20] M.E. Yuksel, Agent-based evacuation modeling with multiple exits using NeuroEvolution of Augmenting Topologies, Adv. Eng. Inf. 35 (November 2016) (2018) 30–55, https://doi.org/10.1016/j.aei.2017.11.003.
[21] R. Zhao, et al., Fuzzy neural network based scenario features extraction and mapping model for crowd evacuation stability analysis, J. Phys. Conf. 1176 (3) (2019), https://doi.org/10.1088/1742-6596/1176/3/032022.
[22] A. Alahi, et al., Social LSTM: human trajectory prediction in crowded spaces, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem, 2016, pp. 961–971, https://doi.org/10.1109/CVPR.2016.110.
[23] I. Hasan, et al., MX-LSTM: mixing tracklets and vislets to jointly forecast trajectories and head poses, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, pp. 6067–6076, https://doi.org/10.1109/CVPR.2018.00635.
[24] Manh, H., & Alaghband, G. (2018). Scene-lstm: A model for human trajectory prediction. *arXiv preprint arXiv:1808.04018*.
[25] H. Xue, D.Q. Huynh, M. Reynolds, SS-LSTM: a hierarchical LSTM model for pedestrian trajectory prediction, in: Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018, 2018, pp. 1186–1194, https://doi.org/10.1109/WACV.2018.00135, 2018-Janua.
[26] G. Hu, et al., When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition, in: Proceedings of the IEEE International Conference on Computer Vision, 2016, pp. 384–392, https://doi.org/10.1109/ICCVW.2015.58.
[27] M. Szarvas, et al., Pedestrian detection with convolutional neural networks, in: IEEE Intelligent Vehicles Symposium, Proceedings, 2005, 2005, pp. 224–229, https://doi.org/10.1109/IVS.2005.1505106.
[28] L. Zhang, et al., Is faster R-CNN doing well for pedestrian detection? Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9906 LNCS (2016) 443–457, https://doi.org/10.1007/978-3-319-46475-6_28.
[29] J. Li, et al., Scale-aware fast R-CNN for pedestrian detection, IEEE Trans. Multimed. 20 (4) (2018) 985–996, https://doi.org/10.1109/TMM.2017.2759508.
[30] P.K.Y. Wong, et al., Recognition of pedestrian trajectories and attributes with computer vision and deep learning techniques, Adv. Eng. Inf. 49 (January) (2021) 101356, https://doi.org/10.1016/j.aei.2021.101356.
[31] L. Boominathan, S.S.S. Kruthiventi, R. Venkatesh Babu, CrowdNet: a deep convolutional network for dense crowd counting, in: MM 2016 - Proceedings of the 2016 ACM Multimedia Conference, 2016, pp. 640–644, https://doi.org/10.1145/2964284.2967300.
[32] J. Gao, et al., 'Cˆ3 Framework: An Open-source PyTorch Code for Crowd Counting' (2019) 3–6. Available at: http://arxiv.org/abs/1907.02724.
[33] F. Camara, et al., Pedestrian models for autonomous driving Part II: high-level models of human behavior, IEEE Trans. Intell. Transport. Syst. 22 (9) (2021) 5453–5472, https://doi.org/10.1109/TITS.2020.3006767.
[34] A. Rudenko, et al., Human motion trajectory prediction: a survey, Int. J. Robot Res. (2020), https://doi.org/10.1177/0278364920917446.
[35] D.A. Zhogolev, V.B. Volkov, B.K. Bunyatyan, Calculation of the electric characteristics of aqua complexes of ions of the 3d transition metals in square and cubic coordinates by the extended Hückel method, Theor. Exp. Chem. (1976) 461–463, https://doi.org/10.1007/BF00523873.
[36] P. Kothari, S. Kreiss, A. Alahi, Human trajectory forecasting in crowds: a deep learning perspective, IEEE Trans. Intell. Transport. Syst. 23 (7) (2022) 7386–7400, https://doi.org/10.1109/TITS.2021.3069362.
[37] A. Vemula, K. Muelling, J. Oh, Social attention: modeling attention in human crowds, in: Proceedings - IEEE International Conference on Robotics and Automation, 2018, pp. 4601–4607, https://doi.org/10.1109/ICRA.2018.8460504.
[38] A. Gupta, et al., Social GAN: socially acceptable trajectories with generative adversarial networks, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, pp. 2255–2264, https://doi.org/10.1109/CVPR.2018.00240.
[39] B. Cheng, et al., Pedestrian trajectory prediction via the Social-Grid LSTM model, J. Eng. 2018 (16) (2018) 1468–1474, https://doi.org/10.1049/joe.2018.8316.
[40] N. Nikhil, B.T. Morris, Convolutional neural network for trajectory prediction, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11131 LNCS (2019) 186–196, https://doi.org/10.1007/978-3-030-11015-4_16.
[41] N. Bisagno, B. Zhang, N. Conci, Group LSTM: group trajectory prediction in crowded scenarios, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11131 LNCS (2019) 213–225, https://doi.org/10.1007/978-3-030-11015-4_18.
[42] Amir Sadeghian, et al., SoPhie: an attentive GAN for predicting paths compliant to social and physical constraints, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019-June, 2019, pp. 1349–1358, https://doi.org/10.1109/CVPR.2019.00144.
[43] Pu Zhang, et al., SR-LSTM: state refinement for lstm towards pedestrian trajectory prediction, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019-June(March), 2019, pp. 12077–12086, https://doi.org/10.1109/CVPR.2019.01236.
[44] Y. Zhu, et al., StarNet: pedestrian trajectory prediction using deep neural network in star topology, IEEE International Conference on Intelligent Robots and Systems (2019) 8075–8080, https://doi.org/10.1109/IROS40897.2019.8967811.
[45] Z. Pei, et al., Human trajectory prediction in crowded scene using social-affinity Long Short-Term Memory, Pattern Recogn. 93 (2019) 273–282, https://doi.org/10.1016/j.patcog.2019.04.025.
[46] V. Kosaraju, et al., Social-BiGAT: multimodal trajectory forecasting using bicycle-GAN and graph attention networks, Adv. Neural Inf. Process. Syst. 32 (2019) 1–10.
[47] J. Cai, et al., A context-augmented deep learning approach for worker trajectory prediction on unstructured and dynamic construction sites, Adv. Eng. Inf. 46 (September) (2020) 101173, https://doi.org/10.1016/j.aei.2020.101173.
[48] A. Mohamed, et al., SociAl-STGCNN: a social spatio-temporal graph convolutional neural network for human trajectory prediction, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020, pp. 14412–14420, https://doi.org/10.1109/CVPR42600.2020.01443.
[49] Dan, X. (2020). Spatial-temporal block and LSTM network for pedestrian trajectories prediction. *arXiv preprint arXiv:2009.10468*.
[50] J. Yang, et al., PTPGC: pedestrian trajectory prediction by graph attention network with ConvLSTM, Robot. Autonom. Syst. 148 (2022) 103931, https://doi.org/10.1016/j.robot.2021.103931.
[51] L. Shi, et al., Social interpretable tree for pedestrian trajectory prediction, Proceedings of the 36th AAAI Conference on Artificial Intelligence 36 (2022) 2235–2243, https://doi.org/10.1609/aaai.v36i2.20121. *AAAI 2022*.
[52] J. Duan, et al., Complementary attention gated network for pedestrian trajectory prediction, Proceedings of the 36th AAAI Conference on Artificial Intelligence, AAAI 2022 36 (2022) 542–550, https://doi.org/10.1609/aaai.v36i1.19933.

[53] S. Yi, H. Li, X. Wang, Understanding pedestrian behaviors from stationary crowd groups, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June, 2015, pp. 488–496, https://doi.org/10.1109/CVPR.2015.7298971.

[54] S. Pellegrini, et al., You'll never walk alone: modeling social behavior for multi-target tracking, in: Proceedings of the IEEE International Conference on Computer Vision, 2009, pp. 261–268, https://doi.org/10.1109/ICCV.2009.5459260. September 2009.

[55] A. Lerner, Y. Chrysanthou, D. Lischinski, Crowds by example, Comput. Graph. Forum 26 (3) (2007) 655–664, https://doi.org/10.1111/j.1467-8659.2007.01089.x.

[56] J. Shao, C.C. Loy, X. Wang, Scene-independent group profiling in crowd, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014, pp. 2227–2234, https://doi.org/10.1109/CVPR.2014.285.

[57] B. Zhou, X. Wang, X. Tang, Understanding collective crowd behaviors: learning a Mixture model of Dynamic pedestrian-Agents, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2012, pp. 2871–2878, https://doi.org/10.1109/CVPR.2012.6248013.

[58] B. Benfold, StableMulti-TargetTrackinginReal-TimeSurveillanceVideo_cvpr2011.pdf, IEEE Conference on Computer Vision and Pattern Recognition, 2011.

[59] N. Lee, et al., DESIRE: distant future prediction in dynamic scenes with interacting agents, in: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017, pp. 2165–2174, https://doi.org/10.1109/CVPR.2017.233, 2017-Janua.

[60] J. Shao, C.C. Loy, X. Wang, Learning scene-independent group descriptors for crowd understanding, IEEE Trans. Circ. Syst. Video Technol. 27 (6) (2017) 1290–1303, https://doi.org/10.1109/TCSVT.2016.2539878.

[61] A. Robicquet, et al., Learning social etiquette: human trajectory understanding in crowded scenes, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9912 LNCS, 2016, pp. 549–565, https://doi.org/10.1007/978-3-319-46484-8_33.

[62] X. Song, et al., Pedestrian trajectory prediction based on deep convolutional LSTM network, IEEE Trans. Intell. Transport. Syst. 22 (6) (2021) 3285–3302, https://doi.org/10.1109/TITS.2020.2981118.

[63] H. Minoura, et al., Path predictions using object attributes and semantic environment, VISIGRAPP 2019 - Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications 5 (Visigrapp) (2019) 19–26, https://doi.org/10.5220/0007297500190026.

[64] H. Xue, D.Q. Huynh, M. Reynolds, Location-velocity attention for pedestrian trajectory prediction, in: Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019, 2019, pp. 2038–2047, https://doi.org/10.1109/WACV.2019.00221.

[65] M. Wang, et al., Unsupervised pedestrian trajectory prediction with graph neural networks, in: Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, 2019-Novem, 2019, pp. 832–839, https://doi.org/10.1109/ICTAI.2019.00119.

[66] J. Chung, et al., 'Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling' (2014) 1–9. Available at: http://arxiv.org/abs/1412.3555.

[67] A. Santoro, et al., A simple neural network module for relational reasoning, Adv. Neural Inf. Process. Syst. (2017) 4968–4977, 2017-Decem.

[68] M. Goldhammer, et al., Camera based pedestrian path prediction by means of polynomial least-squares approximation and multilayer perceptron neural networks, in: IntelliSys 2015 - Proceedings of 2015 SAI Intelligent Systems Conference, 2015, pp. 390–399, https://doi.org/10.1109/IntelliSys.2015.7361171.

[69] Y. Ma, E.W.M. Lee, R.K.K. Yuen, An artificial intelligence-based approach for simulating pedestrian movement, IEEE Trans. Intell. Transport. Syst. 17 (11) (2016) 3159–3170, https://doi.org/10.1109/TITS.2016.2542843.

[70] S. Pengyuan, A more realistic simulation of pedestrian based on cellular automata, in: OSSC-2009 - Proceedings of 2009 IEEE International Workshop on Open-Source Software for Scientific Computation, 2009, pp. 24–29, https://doi.org/10.1109/OSSC.2009.5416792.

[71] X. Song, et al., A data-driven neural network approach to simulate pedestrian movement, Phys. Stat. Mech. Appl. 509 (2018) 827–844, https://doi.org/10.1016/j.physa.2018.06.045.

[72] X. Zhao, et al., Artificial neural network based modeling on unidirectional and bidirectional pedestrian flow at straight corridors, Phys. Stat. Mech. Appl. 547 (2020) 123825, https://doi.org/10.1016/j.physa.2019.123825.

[73] S. Sharma, et al., Intelligent agents in a goal finding application for homeland security, Conference Proceedings - IEEE SOUTHEASTCON (2012) 9–13, https://doi.org/10.1109/SECon.2012.6197067.

[74] J.K.K. Yuen, et al., An intelligence-based optimization model of passenger flow in a transportation station, IEEE Trans. Intell. Transport. Syst. 14 (3) (2013) 1290–1300, https://doi.org/10.1109/TITS.2013.2259482.

[75] Y. Ma, et al., An intelligence-based approach for prediction of microscopic pedestrian walking behavior, IEEE Trans. Intell. Transport. Syst. 20 (10) (2019) 3964–3980, https://doi.org/10.1109/TITS.2019.2931892.