Neural network model for predicting variation in walking dynamics of pedestrians in social groups

Sun, Shi; Sun, Cheng; Duives, Dorine C.; Hoogendoorn, Serge P.

DOI
10.1007/s11116-021-10263-8

Publication date
2022

Document Version
Final published version

Published in
Transportation

Citation (APA)
Sun, S., Sun, C., Duives, D. C., & Hoogendoorn, S. P. (2022). Neural network model for predicting variation in walking dynamics of pedestrians in social groups. *Transportation*, *50*(3), 837-868. https://doi.org/10.1007/s11116-021-10263-8

Transportation
https://doi.org/10.1007/s11116-021-10263-8

Neural network model for predicting variation in walking dynamics of pedestrians in social groups

Shi Sun[1] · **Cheng Sun**[1] · **Dorine C. Duives**[2] · **Serge P. Hoogendoorn**[2]

Accepted: 31 December 2021

Abstract
Pedestrian spaces are increasingly becoming popular locations for shopping, recreation, festivities, and other social activities. Therefore, an improved understanding of the factors that make walking environments enjoyable and safe is essential. Most existing studies focus on modelling walking behaviours of individual pedestrians. However, most people participate in these activities as parts of social groups. Although the movement and choice behaviours of pedestrians in social groups differ from those of individuals, a model featuring group movements has not been developed. This study uses neural networks to analyse the effects of variables influencing pedestrian movements of social groups and predict the variation in walking dynamics. A top-view video was used to extract the trajectories of pedestrian groups. After identifying the social groups in a crowd, the movement characteristics, pedestrian–environment interaction, inter-pedestrian interaction, intra-group relationship, and inter-group relationship of all group members were calculated and considered in the model. After a variable selection process using neural networks, a neural network model was developed featuring variables that are strongly related to the lateral or longitudinal changes in the individual's walking speed. The current movement condition, presence of obstacles nearby, impending collisions, current position and velocity of other group members, and following behaviour were found to impact a pedestrian's walking dynamics. The proposed model can predict the pedestrian density and distribution according to a space function, contributing to better crowd management and efficient design and renovation of pedestrian spaces. Furthermore, the variable selection method can optimise and simplify other pedestrian behaviour prediction models.

Keywords Pedestrian · Social group · Neural network · Group dynamics · Unmanned aerial vehicle (UAV)

✉ Shi Sun
sunshi_1989@163.com

Extended author information available on the last page of the article

Published online: 13 February 2022

Introduction

People enjoy spending their leisure time in public urban spaces, such as shopping malls, game parlours, concerts, and festivals. Many places where people undertake such activities are pedestrian-only spaces because they offer a favourable and safe environment for pedestrians. These public urban spaces have attracted increased attention over time, resulting in overcrowding, which can diminish the pedestrian experience of these spaces and even affect the physical safety owing to crowd incidents.

Pedestrian flow models can be used to predict the walking dynamics of a crowd. In general, pedestrian flow models are divided into microscopic and macroscopic models (Duives et al. 2013). Microscopic models are most often used to simulate a pedestrian's operational movement dynamics. The social force model is one of the most frequently used microscopic models; it considers both physical and motivational factors when determining the next step of each individual agent in the simulated environment. According to this model, the pedestrian walking behaviour is influenced by the goal and desired velocity of the individual, physical boundaries of spaces (e.g. walls, obstacles), and attractive/repulsive effects of objects, spaces, and other agents in the surrounding environment of an agent. In the original version of the social force (SF) model proposed by Helbing and Molnár (1995), the agents were assumed to have no social attachment to other pedestrians.

In recent years, several alterations have been suggested to the SF model to enhance its realism. For instance, Zacharias (2001) found that the gathering of pedestrians might increase the attractiveness of certain locations, and Qu et al. (2018) established that overcrowding creates unfavourable feelings among people and consequently repulse them. Koh and Wong (2013) found that commercial activities and crowdedness are both important factors that divert pedestrians from the preferred path. Furthermore, Xiao et al. (2016) established that the distribution of pedestrian density is an important factor influencing the pedestrian experience in a pedestrian environment.

Moreover, the behaviour of pedestrians moving in groups, known as group dynamics, was found to be an important factor influencing the walking behaviour of individual pedestrians (Lu et al. 2014), crowd movement dynamics (Duives et al. 2014), and distribution of pedestrians over a space (Wang et al. 2013). Considerable research has been conducted on the statistical properties of group dynamics (e.g. James 1953; Bakeman and Beck 1974). For instance, observations on pedestrian group size showed that the sizes of social groups were mostly less than or equal to 6 (Ge et al. 2012; Li et al. 2015; Feng and Li 2016). Additionally, Moussaïd et al. (2010) illustrated that the group size influences the walking speed and formation of a social group. Zanlungo et al. (2014) reported that small pedestrian groups generally exhibit similar constellations.

A common feature of all microscopic pedestrian models featuring group dynamics is that small explicit additions to the model structure have been made to account for the influence of the group. An additional force with a predetermined mathematical formulation is usually added to the array of 'social' forces. Essentially, the estimation procedure comprises the calibration of the parameters of this explicit new force (i.e. reaction strength and radius). However, it should be investigated whether this explicit manner of accounting for the interactions with group members as well as non-group members comprehensively accounts for all influences of the group on the movement dynamics of the group members.

In contrast to these previous attempts, the present study adopts a different approach to the modelling of pedestrian group dynamics. The objective of this study is to derive a neural network model incorporating a comprehensive set of input variables that influence a

pedestrian's movement behaviour in social groups. For this purpose, this study estimates an array of 'simple' artificial neural network models that feature the movement characteristics to evaluate the effect of variables on the prediction accuracy and obtain the best set of input variables. This new model accounts for a wide range of influences, such as the movement characteristics of the agent, pedestrian–environment interaction, inter-pedestrian interaction, intra-group relationship, and inter-group relationship.

The remainder of this paper is structured as follows. "Background" section presents a brief overview of the previous studies on pedestrian movement prediction models. "Methodology" section comprehensively discusses the research methodology. "Case study" section presents a case study. "Results" section presents the results of selecting the related variables and running the neural network model. Finally, "Conclusion" and "Applications and future work" sections, respectively, present the main conclusions of this research and the potential applications of this model to future research and the design of public walking spaces.

Background

Previous studies have attempted to model walking and group dynamics using microscopic pedestrian simulation models and artificial intelligence (AI) models. This section reviews the literature intensively. "Explicit modelling of group dynamics" section discusses the current state-of-the-art pertaining to the modelling of group dynamics. Furthermore, "Machine-learning modelling of walking behaviour (and group dynamics)" section elaborates the use of AI models to model pedestrian movement behaviour.

Explicit modelling of group dynamics

The studies featuring the modelling of group dynamics can be classified based on the foundation model that they enhanced. Three foundation models are used, namely, social force (SF) model, cellular automaton (CA) model, and agent-based model (ABM). The studies that have made specific adaptations to incorporate group dynamics are discussed below.

The SF model is a continuous model that describes the influence of the desired velocity, presence of other pedestrians, physical borders, and attractiveness of nearby locations on the pedestrian behaviour as a force. Several researchers have enhanced the original SF model of Helbing and Molnár (1995) to incorporate group dynamics. Moussaïd et al. (2010) attempted to incorporate head rotation, attraction effect, and repulsive effect in social groups into the SF model. Furthermore, Zanlungo et al. (2014) used interaction with group members and the inter-group collision avoidance behaviour to describe the group dynamics. Recently, Zhang et al. (2018) attempted to incorporate the leader–follower behaviour in an inter-group relationship and the similarity of position and velocity in an intra-group relationship into the SF model. This SF model is capable of describing the interactions, but the force having the dominant influence could not be clarified.

Furthermore, the CA model divides a space into grids of cells and simulates the movement of individuals between cells in discrete time intervals. This model has been improved to simulate group behaviour in crowds. For instance, Köster et al. (2011) incorporated social groups into the CA model and found that group behaviour influences the crowd movement; several group behaviours were validated experimentally. Recent studies have also incorporated intra-group structure (You et al. 2016) and leader–follower behaviour

(Lu et al. 2017) into the CA model. CA models require less computation and effectively describe the existing theories in a discrete space; however, they are incapable of examining and updating the theory of walking behaviour.

On the other hand, the ABM is a 'bottom-up' model, which allows agents to move in a continuous space instead of grids; it also allows setting movement rules without the restriction of being driven by a 'force'. Qiu and Hu (2010) incorporated an intra-group structure and inter-group relation into the ABM and found that the group size, intra-group structure, and inter-group following behaviour affect the crowd dynamics. Later, researchers attempted to incorporate interactions between sub-groups (Fu et al. 2014), panic case (Wang et al. 2015), and social group speed coordination (Kiefer et al. 2017) into the ABM. In the ABM, the rule of movement is predetermined, making it feasible to execute a variety of behaviours; however, similar to the CA model, the predetermined rule follows the previously studied theory, and new influencing factors cannot be incorporated.

In summary, attempts have been made in an unorganised manner to simulate group behaviour by means of microscopic pedestrian models. Previous studies showed the influence of group dynamics on the pedestrian walking behaviour by adding variables describing group dynamics, but the relative importance of variables describing group dynamics and other variables were not directly compared in a comprehensive model. Therefore, further fundamental research is required to determine the exact factors that influence group movement behaviours.

Machine-learning modelling of walking behaviour (and group dynamics)

AI models allow the implicit modelling of various movement dynamics (Chella et al. 2000). In other words, machine learning models can learn a relationship between input and output variables, while allowing much flexibility in specifying prior assumptions on this relationship. These model types can also be used to study pedestrian movement behaviour. This section reviews the three most promising AI modelling techniques that have already been used by other researchers to model walking behaviour, namely, the Markov model, neural network, and long-short term memory (LSTM) model.

Markov model

A Markov chain is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. A Markov model can be effectively used to study a discrete system and can be applied to the spatial sequential choice. Models such as the hidden Markov model (HMM), relational Markov networks (RMN), and Markov decision process (MDP) use the Markov chain concept to predict pedestrian behaviour. Ashbrook and Starner (2003) incorporated previous locations into a Markov model to predict the movement of people. By considering individual characteristics, Gambs et al. (2010) used a Markov model to predict the next destination of a person based on the historical information of the previously visited spaces and the probability distribution of transitions between states. Furthermore, Gambs et al. (2012) included several historical locations of the concerned person to improve the precision of prediction.

To study the movement and distribution of pedestrians in a public area, a model that can be applied to predict the instantaneous movement behaviour is required to simulate pedestrian walking dynamics. Nascimento et al. (2010) proposed a two-layer HMM to predict

the pedestrian movement dynamics based on the historical trajectories. The lower-level model features stopping and moving to the north, west, south, or east, and the upper-level model switches the action mode of the lower level according to the type of activity. Burkert and Bamler (2012) incorporated pedestrian interaction by calculating four motion features between pedestrians, and a weight was assigned for the interaction based on the distance between the pedestrians. This model considered the interaction between neighbouring pedestrians, and it could make sense for determining social groups with two pedestrians. The Markov models verified the importance of historical states in predicting pedestrian dynamics. Because Markov models aim to predict discrete states, the amount of input variables required to build a continuous prediction model is extremely large, and it is difficult to calibrate the model to compare the effects of different variables.

Neural network

The neural network considered in the present study excludes deep learning, which distinguishes its concept in this study from the LSTM model. A neural network is a series of algorithms that endeavours to recognise the underlying relationships in a set of data through a process similar to the operation of the human brain. Zheng et al. (2002) combined the SF model and a neural network and found that an appropriate proportion of impatient pedestrians in a crowd could improve the crowd movement and passage time. First, a neural network was used to build a choice model of pedestrian behaviour (Ottomanelli et al. 2010; Zainuddin and Lim 2012; Yuen et al. 2014). Later, Yi et al. (2016) used the previous displacement of the pedestrian to predict the future movement. Song et al. (2018) compared the neural network model and SF model in several scenarios and found that the former exhibits better performance in predicting the pedestrian movement than the latter. In general, previous neural network models used historical information and the surrounding condition as inputs and considered repulsion from the walking environment.

Because pedestrian movement is also influenced by the attraction from group members, a prediction model for walking behaviour influenced by social groups is yet to be constructed. Wang et al. (2019) proposed a method to select the input variables from a group of behavioural and environmental factors that could be used to predict the variation in walking speed to simplify the prediction model. However, that study focused on a choice model of acceleration, deceleration, and maintenance of the current speed, which do not quantify the change in speed or consider the change in direction. Thus far, there have been no reports on a neural network model that considers a variety of variables to predict the trajectory and walking dynamics and is combined with a method to filter the variables in this model.

LSTM model

The LSTM model is a type of recurrent neural network model that has recently been used to predict pedestrian trajectories. It is capable of considering all historical information and discarding irrelevant data, enabling it to automatically learn the current and historical states of a pedestrian and predict the future positions. Alahi et al. (2016) used Soical-LSTM networks to predict the pedestrian movement by considering the previous movement condition and spatial neighbours in the model. It was found that the LSTM model could predict pedestrian movement, but the environmental factors were not included in this model. Pfeiffer et al. (2018), Bartoli et al. (2018), and Xue et al. (2018) incorporated

the environmental borders and obstacles in the LSTM model and improved the prediction accuracy. Later, other factors, such as the attention model (Fernando et al. 2018) and spatial affinity between pedestrians (Xu et al. 2018), were integrated with the LSTM model to further improve the accuracy. Recently, Bisagno et al. (2019) considered social groups in the LSTM model by treating them as an entity and predicted the movement dynamics by considering only the influence of the inter-group relationship. In that study, many other factors such as leader–follower interaction and intra-group interaction were neglected. Fernando et al. (2019) used an LSTM model to predict the pedestrian's operative movement behaviour and attempted to derive the influence of group behaviour from this movement prediction model. That study found an approach to extract pedestrian social groups in the crowd by identifying group behaviour, but the influence of group members on the walking dynamics was not quantified. Generally, in the LSTM model of predicting pedestrian dynamics, the influence from walking environment and other pedestrians are represented by their hidden states, and the type of interaction is not specified. Moreover, a prediction model for pedestrian movement considering social group behaviour has not been explicitly elucidated.

In general, the Markov model has logic to predict the pedestrian dynamics, but it cannot cope with the calibration of many variables needed to be incorporated in a continuous space. The neural network and LSTM model are not limited by multicollinearity between variables since their aim is to achieve the highest fit without identifying functional relationships of individual variables (Lakes et al. 2009). Because the influencing factors of walking dynamics may have multicollinearity, and most other models should not take the influencing factors with multicollinearity as input variables simultaneously, which will lead to the exclusion of some influencing factors, neural network and LSTM model can establish a comprehensive model. The simplicity of neural networks may impair their accuracy in some cases; however, because of their simplicity, they have the potential to be comprehensive and comprehensible. The LSTM model is capable of making accurate predictions; however, because of its complexity, no existing method has been sufficiently validated to determine the relative importance of influencing factors represented by hidden states. Regardless of the type of model used for prediction, a proper set of input variables is always a key factor. A neural network provides the opportunity to use a variety of variables as inputs simultaneously and calculate the contribution of each variable to the walking dynamics to obtain a prediction model with the set of most influential input factors.

Methodology

The present study attempts to establish a comprehensive set of factors that are jointly required to model grouping behaviour. Thus, a model type is required that can be interpreted mathematically and applicable in variable selection. Simultaneously, we intend to step away from the explicit modelling of the variables and the relations between them. A neural network is suitable for this purpose. This section presents a method to model pedestrian group behaviour implicitly by means of a neural network.

First, the general structure of the neural network is presented. Accordingly, the variables that feed the model are introduced; then, the variable selection process is detailed. Finally, the parameter analysis method used to handle the stochasticity in the model calibration process is discussed.

Design of neural network model

As seen in Fig. 1, the neural network model in this study is a three-layer network comprising an input layer, a hidden layer, and an output layer, because the neural network with one hidden layer can be more comprehensible (Heaton et al. 2017). The input layer contains one or more input variables used to predict the output. The output layer is the acceleration along the x-axis (N_x) and y-axis (N_y) in the next time interval (0.5 s). The number of neurons in the hidden layer correlates to those in the input layer. The activation function between input layer neurons and hidden layer neurons is tangent sigmoid function, and the activation function between hidden layer neurons and output layer neurons is linear transfer function. The tangent sigmoid function is defined as

$$y = \frac{2}{1 + e^{-2x}} - 1, \tag{1}$$

where x denotes the input of hidden neurons, and y represents the output of hidden neurons.

The cases are divided into two sets: training and testing. The training set is used to train the model (determine the coefficients for the multinomial logit model and parameters for the machine learning model). The testing set is used to test the performance of the trained model. The predicted results from the testing set are compared with the actual experimental data to determine the final prediction accuracy. In this study, the training and testing groups are randomly selected from the dataset. A limited number of groups featuring large group sizes are present in the dataset, which might cause an imbalance in the test of the model. Thus, all the groups are divided into two parts to avoid imbalance between the training

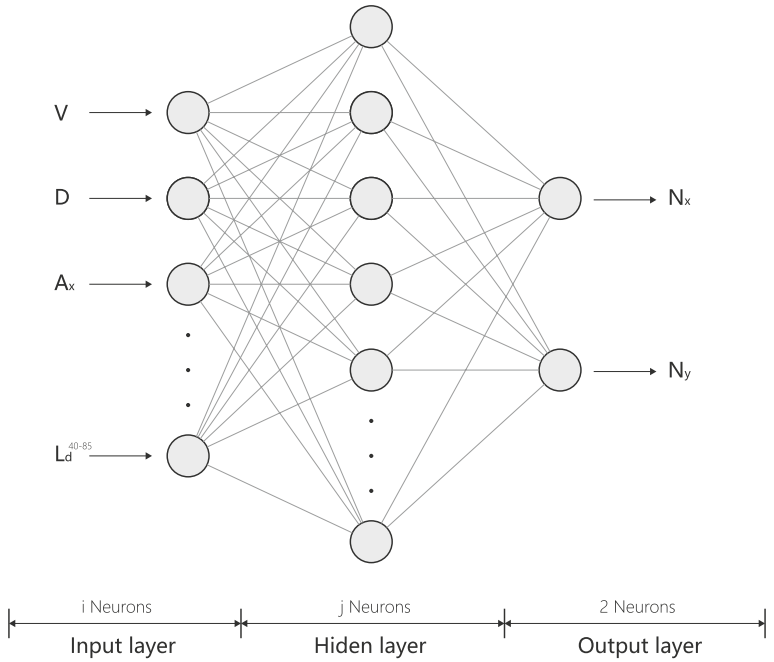


Fig. 1 Architecture of neural network used in this study

and testing data. Groups with more than three members are classified as large groups, and the others are classified as small groups. In the process of dividing the data, 80% of small groups and 80% of large groups are selected to be used for training, and the remaining 20% of the data featuring both types of groups are used for testing.

Identification of candidate input variables

A pedestrian can be affected by his/her current and historical states (Alahi et al. 2016). In this study, the historical state is described by the velocity variation in the previous time interval. For the current state, according to the SF model, the pedestrian walking dynamics is influenced by the preferred velocity, repulsive force from the borders or other pedestrians, and influence of social groups (Helbing et al. 2001). The influence of social groups can be further classified into intra-group relationships and inter-group relationships (Qiu and Hu 2010). As shown in Fig. 2, the input variables considered in the model can be classified into five groups: movement characteristics, pedestrian–environment interaction, inter-personal interaction, intra-group relationship, and inter-group relationship. In order to organise the complex environment influence (non-fixed number of obstacles, neighbours, and leaders) as a few input variables to be used in the neural network, basic rules are used to convert these influences into variables.

The first group of input variables describes the movement characteristics. In most microscopic models, these comprise the walking speed, direction, and acceleration. Thus, these variables are included in the set of candidate variables for this neural network. In

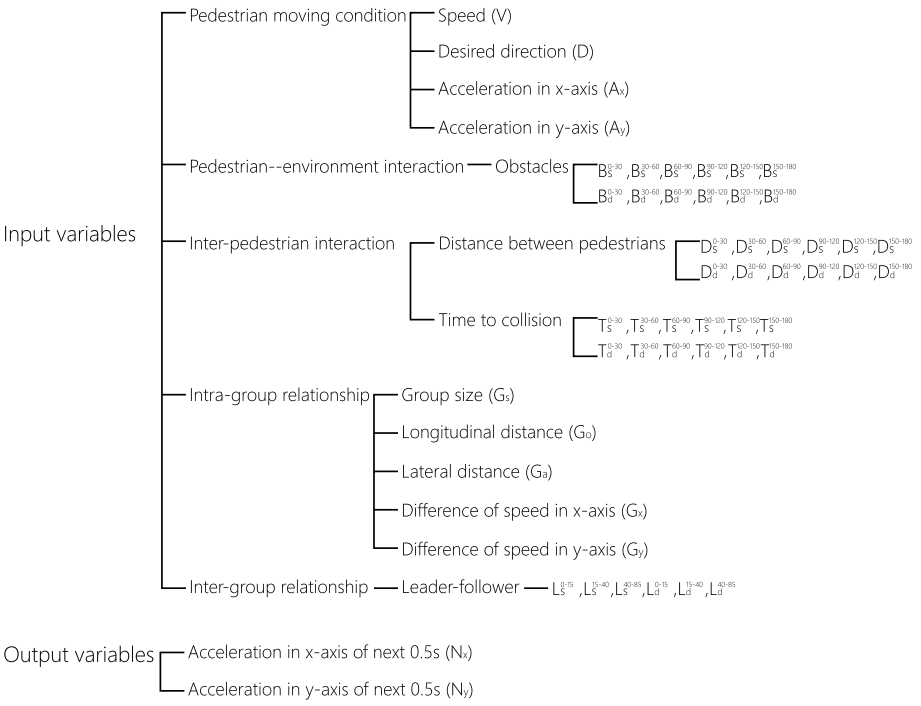


Fig. 2 Structure of the input and output variables in the model

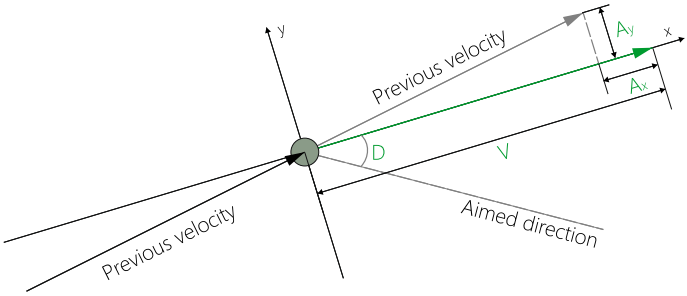


Fig. 3 Variables describing the movement characteristics

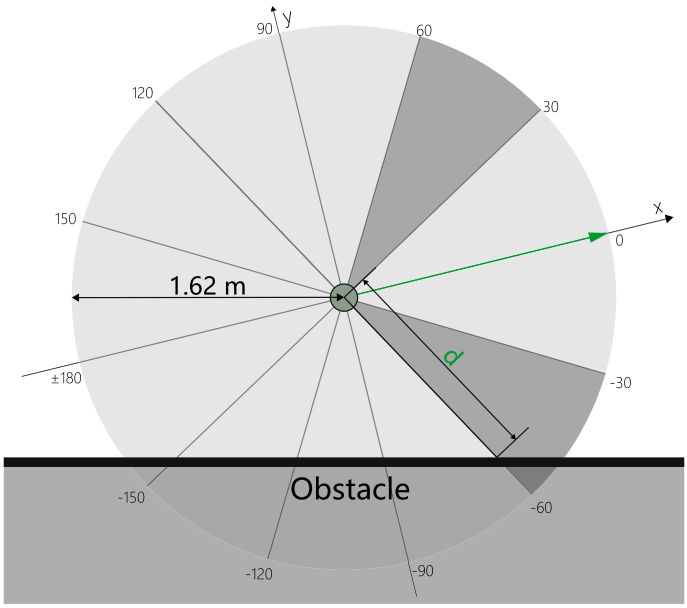


Fig. 4 Interaction between pedestrians and obstacles in the environment

this research, this group of variables comprises the current speed (V), desired direction (D), acceleration along the x-axis (A_x), and acceleration along the y-axis (A_y), as shown in Fig. 3. Because the walking dynamics in a short time interval is influenced by the body swing, and a long time interval fails to describe certain details of the walking process, the time interval is set to 0.5 s. V is calculated as the distance of movement in the previous 0.5 s divided by the time interval (0.5 s). The intended direction is the direction extending from the current position to the final position in the area of focus, and D is the angle between the current walking direction and the intended direction. A_x and A_y denote the difference between the current velocity and the velocity 0.5 s earlier along the x-axis and y-axis of the current coordinate system divided by the time interval (0.5 s), respectively.

The second group of variables features the interaction between the pedestrians and obstacles in the environment, as shown in Fig. 4. Pedestrians maintain a safe distance from other elements in the environment to avoid the risk of injury (Helbing and Molnár

1995). Therefore, in this study, walls, handrails, steps, and slopes are treated as obstacles. The force from the obstacles within a distance of 1.62 m (Johansson et al. 2007) from a specific pedestrian is considered. Borders such as steps and slopes, which can be navigated, are generally considered obstacles. However, in the case of a pedestrian about to leave the region of interest from a border, the border is not treated as an obstacle when the distance between the current pedestrian position and the position on the border from where the specific pedestrian will exit the region is less than 4 m (Fajen and Warren 2003). Based on the angle of interaction, the obstacles are divided into 12 groups, namely, 0°–30°; 0° to − 30°; 30°–60°; − 30° to − 60°; 60°–90°; − 60° to − 90°; 90°–120°; − 90° to − 120°; 120°–150°; − 120° to − 150°; 150°–180°; and − 150° to − 180°. In this study, the symmetrical parts of the region around the current walking direction, such as 30°–60° and − 30° to − 60°, are treated as opposite positions. Furthermore, the sum of and difference in the forces at the opposite positions are calculated and considered as the input variables of this group. For the force in the region from α to β, $B_s^{\alpha,\beta}$ is the sum of forces generated from the obstacles, and $B_d^{\alpha,\beta}$ is the difference in the forces. The force generated by an obstacle decays exponentially with the distance from the edge of the specific pedestrian (Johansson et al. 2007). Thus, if the radii of pedestrians are considered as 0.3 m, the force is calculated using Equation 2.

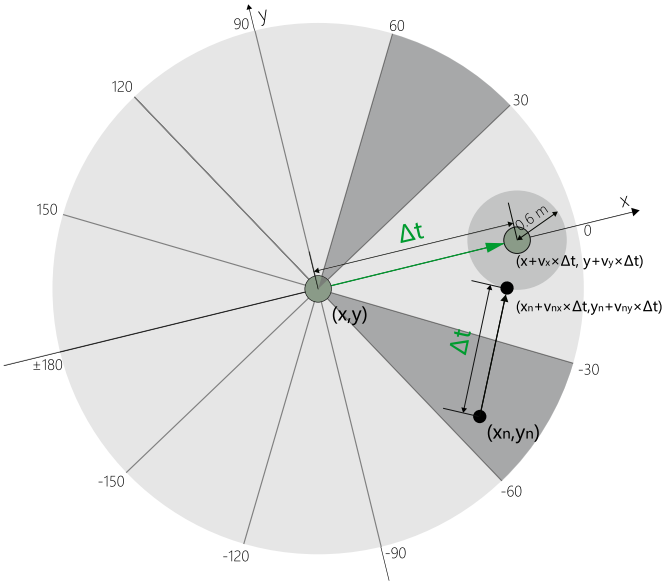
$$F(n) = e^{(0.3-d)} \tag{2}$$

The third group of variables pertains to interpersonal interactions. This group of variables includes the time to collision and the distance between pedestrians. Considering the angle between the current walking direction of a specific pedestrian and the position of other pedestrians, to clarify the influence from specific directions, other pedestrians are separated into 12 groups using the same method as separating obstacles, as shown in Fig. 5. This group of variables are also the sum of and difference in the forces generated from the opposite positions of the walking direction.

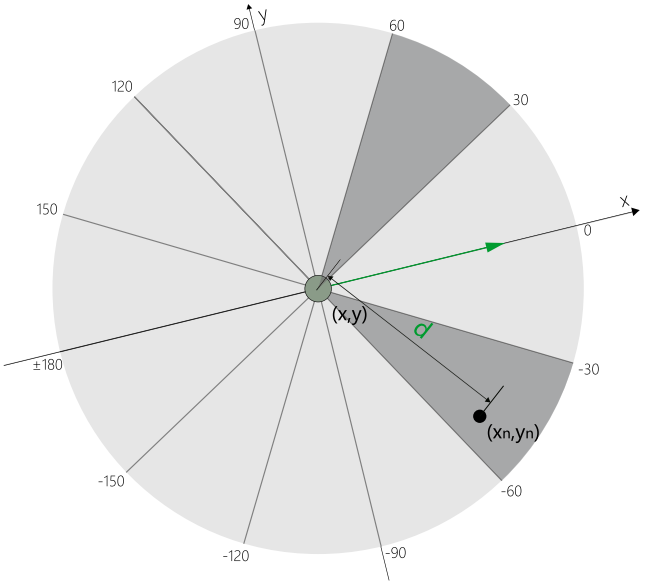
As shown in Fig. 5a, time to collision is the time until an impending collision. A situation where the distance between a specific pedestrian and another pedestrian is less than 0.6 m is regarded as a collision, and the time (Δt) is calculated. The time to collision is also considered as a force. For the force in the region from α to β, $T_s^{\alpha,\beta}$ is the sum of forces generated from the pedestrians, and $T_d^{\alpha,\beta}$ is the difference in the forces. Because the probability of collision decays exponentially with Δt (Festa et al. 2018), the influence of a pedestrian (n) is calculated using Equation 3. If no possible collisions arise from a specific region, the force will be considered as 0.

$$F(n) = e^{-\Delta t}. \tag{3}$$

Other pedestrians can also be treated as obstacles. For the direction from α to β, $D_s^{\alpha,\beta}$ is the sum of forces between α and β, and $D_d^{\alpha,\beta}$ is the difference in the forces between α and β. For another pedestrian (n), the distance from the specific pedestrian (d) is used to calculate the force, as seen in Fig. 5b. The force generated by other pedestrians decays exponentially with the distance between the edges of these two pedestrians (Johansson et al. 2007); thus, if the radii of pedestrians are considered as 0.3 m, the distance between the two pedestrian edges is $d - 0.6$ m, and the distance-dependent force is calculated using Eq. 4. If no other pedestrians appear in a specific region, the distance-dependent force will be considered as 0.



(a) Time to collision with other pedestrians



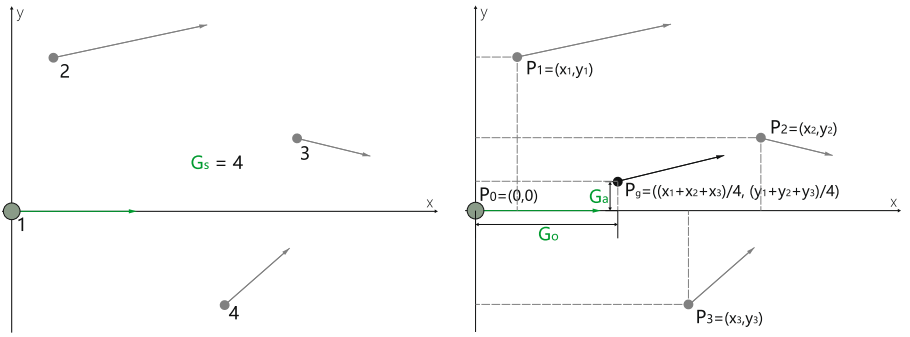
(b) Distance between a specific pedestrian and other pedestrians

Fig. 5 Interpersonal interaction

$$F(n) = e^{(0.6-d)}. \tag{4}$$

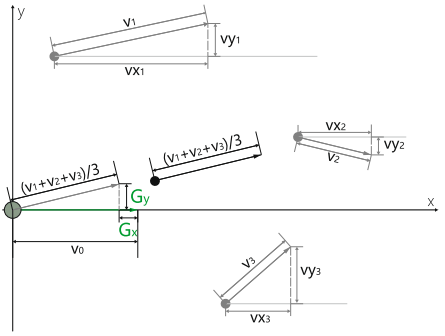
The fourth group of variables comprises intra-group relationships. Aggregation and following are the two actions that group members take to maintain the preferred intra-group structure (Qiu and Hu 2010), and group size influences the preferred structure (Moussaïd et al. 2010). As seen in Fig. 6, this group of variables includes group size, relative positions of the pedestrians within a group, and the difference in velocity between a specific pedestrian and other group members. Group size (G_s) defines the number of pedestrians in the social group. The central position is the mean value of the positions of all group members on the x-axis and y-axis. The longitudinal and lateral distances between the specific pedestrian and the central position on the x-axis and y-axis are G_o and G_a, respectively, which describe the aggregation in the social group. The difference in velocity includes two variables: G_x and G_y. These two variables indicate the difference in the walking speed between the specific pedestrian and the mean values of the other group members on the x-axis and y-axis, which describe the following behaviour in the group.

The fifth group of variables relates to the inter-group interaction; this group of variables includes the leader–follower behaviour, as shown in Fig. 7 (Qiu and Hu 2010). Robin et al. (2009) modelled the leader–follower behaviour and divided the visible region (− 85° to



(a) Group size

(b) Relative position of a specific pedestrian in a social group



(c) Relative velocity of a specific pedestrian with other group members

Fig. 6 Intra-group relationship

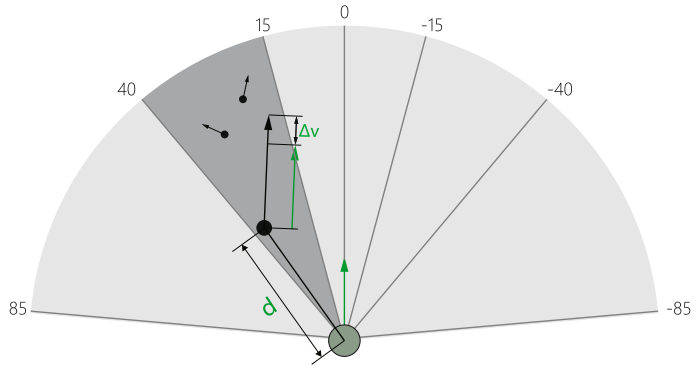


Fig. 7 Relationship between a leader and a specific pedestrian

85°) into a 'central' region (− 15° to 15°), 'side' region (− 45° to − 15° and 15° to 45°), and 'extreme' region (− 85° to − 45° and 45°–85°). In the present study, these regions are separated as left and right regions. The left region comprises the regions 0°–15°, 15°–40°, and 40°–85°, and the right region comprises the regions 0° to − 15°, − 15° to − 40°, and − 40° to − 85°. Pedestrians outside the targeted social group but within the range of 6.25 m from a specific pedestrian and have a walking direction between − 10° and 10° are selected as potential leaders. Among all the potential leaders, the one nearest to the specific pedestrian is the leader at that moment. The variables of leader–follower behaviour are the sum of and difference in the forces generated from opposite positions of the walking direction. For the direction from α to β, $L_s^{\alpha,\beta}$ is the sum of the leader–follower force between α and β, and $L_d^{\alpha,\beta}$ is the difference in the leader–follower force between α and β. The attractive force from the leader is positively correlated with the difference in speed (Δv) and negatively correlated with the distance between the specific pedestrian and the leader (d) (Robin et al. 2009). The variable $\Delta v = v_l - v_p$, where v_l is the speed module of the leader and v_p is the speed module of the specific pedestrian. For a leader (n), the following force is calculated using Eq. 5. Faster leaders cause acceleration, and slower ones induce deceleration, but all leaders attract the specific pedestrian along the y-axis. Because the difference in forces from opposite positions is correlated to the movement dynamics on the y-axis, $|F(n)|$ is regarded as the leader–follower force when calculating $L_d^{\alpha,\beta}$.

$$F(n) = (v_l - v_p)/d. \tag{5}$$

Neural network model selection process

A branch and bound technique is used to derive the neural network model with the best prediction accuracy. The process is initially simple but gains complexity with the progressive addition of variables into the model. Three procedures are used to select the best neural network to describe the overall walking behaviour, including group behaviour, in the dataset. In the first procedure, variables without enough prediction ability are filtered out. In the second procedure, variables with sufficient relative importance are selected. Finally, the number of neurons is determined according to the average displacement error (ADE) of the prediction model, as shown in Fig. 8.

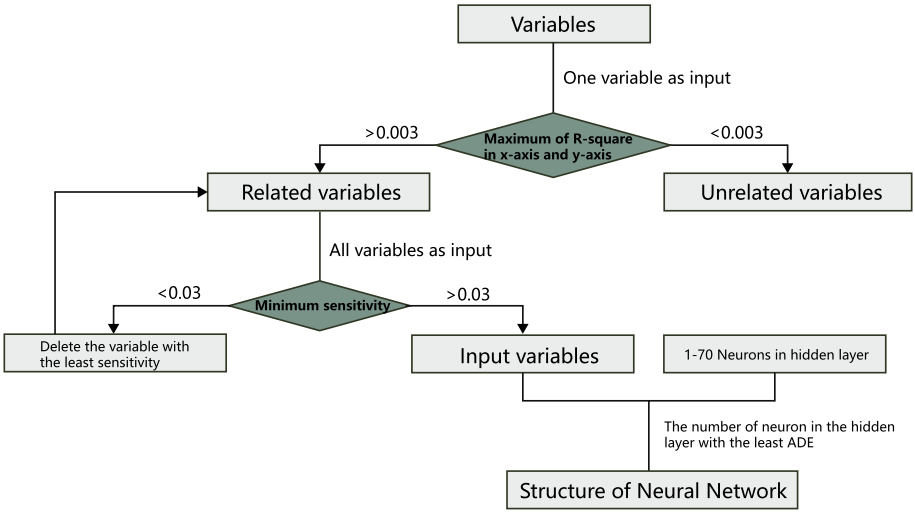


Fig. 8 Process of variable selection

In the first procedure, the ability of each candidate variable to predict the walking behaviour is determined. For this purpose, a neural network model is estimated that uses only one candidate variable as the input at a time. The prediction ability of every input variable is evaluated by R-square, which indicates the prediction ability (Tan et al. 2002) and can be used for both linear and non-linear models (Ržazanskasa et al. 2017; Elmchichi et al. 2020). The R-square is defined as

$$R^2 = \frac{\sum (y - \bar{y})^2 - \sum (y - \hat{y})^2}{\sum (y - \bar{y})^2}, \tag{6}$$

where y denote ground truth, $\bar{y}$ is the mean value of ground truth, and $\hat{y}$ means the predicted value. The R-square closer to 1 indicates better prediction ability. The R-square between the prediction result and ground truth in the testing process is accordingly used to filter the variables with sufficient prediction ability. This procedure involves three steps.

In step 1, the R-square between the prediction result of every variable as input and the ground truth is calculated repeatedly for a fixed number of hidden neurons (n). Accordingly, the mean value of all the previous repetitions is calculated and recorded. Therefore, the results of 'step n' comprise two sets, namely, the mean value of all repetitions up to that attempt in and excluding the current repetition. If the margin of error in the 95% confidence interval of both sets is less than 0.0001 or the number of repetitions exceeds 999, the repetition will terminate and the median R-square of all repetitions will be recorded as Rx_n and Ry_n. In step 2, the process of step 1 repeats with hidden neurons from 1 to 9 for all models, and the median value of Rx_n and Ry_n will be, respectively, selected as Rx and Ry. In step 3, in order to get a comprehensive input variable group, even variables with weak prediction ability should be included in the model. The R-square above 0 can indicate that a specific variable has prediction ability (Barten 1987). However, due to the randomness of the neural network model, the R-square of variables without prediction ability will fluctuate around 0. Therefore, setting 0 as the threshold will not be able to filter out some variables without prediction ability and obtain a stable candidate set. Then, a threshold that can keep a large number of variables with

prediction ability and filter out variables with R-square fluctuating around 0 to obtain a stable candidate set should be used. In this study, variables with R-square larger than 0.003 in *Rx* or *Ry* are selected for inclusion in the candidate set.

In the second procedure, the neural network model is estimated with all the selected variables as inputs, and the relative importance of every input variable is calculated. As a nonlinear model, determining the relative importance of variables is more difficult than in linear regression models, and Garson's method is used in this study, because it was proved that the neural network can identify the most influential input variables from a given variable list by Garson's method (Garson 1991), and recently, the reliability of Garson's method in the variable selection process of three-layer neural network is verified to be better than other widely used methods, such as correlation method and principal component analysis (Papatheocharous and Andreou 2010; Fischer 2015; Yousefi et al. 2018; Liu et al. 2018). According to Garson's method (Garson 1991), for a neural network model with N neurons in the input layer and L neurons in the hidden layer, the relative importance of the ith input variable to the kth output variable (I_{ik}) can be defined as

$$I_{ik} = \frac{\sum_{j=1}^{L} \left(\left| \omega_{ij} v_{jk} \right| / \sum_{r=1}^{N} \left| \omega_{rj} \right| \right)}{\sum_{r=1}^{N} \sum_{j=1}^{L} \left(\left| \omega_{ij} v_{jk} \right| / \sum_{r=1}^{N} \left| \omega_{rj} \right| \right)}, \tag{7}$$

where ω_{ij} is the weight of the ith neuron in the input layer and jth neuron in the hidden layer, and v_{jk} is the weight of the jth neuron in the hidden layer and kth neuron in the output layer. Variables with sufficient relative importance will be selected for inclusion in the neural network model. This procedure involves three steps.

In step 1, the relative importance of every input variable is calculated by Garson's method in the prediction model with 1–9 hidden neurons (n), and the mean value is recorded as the relative importance in this step. In step 2, because the relative importance calculated by Garson's method can reach a stable trend with 999 repetitions (Olden and Jackson 2002), the process followed in step 1 is repeated 999 times, and the mean value of the relative importance in all the previous repetitions is calculated and recorded as I. In step 3, steps 1 and 2 are repeated until the value of I for every input variable is larger than a certain threshold. Because a high threshold can simplify the model, and a low threshold can ensure the higher prediction accuracy, in order to exclude less influential values and maintain a sufficiently comprehensive set of input variables, the threshold is set to 0.03. If the value of I for an input is less than 0.03, the input with the least I will be deleted, and the updated set of the remaining variables will be the input in the next repetition of steps 1 and 2. The variables remaining after all repetitions are the set of candidate input variables.

At the end of the variable selection process, the ADE of the neural network model featuring the selected variables as input is calculated repeatedly with a fixed number of hidden neurons (n), and every repetition is recorded. For all N pedestrians appearing from t_{sta} to t_{fin}, the ADE of each repetition is calculated as

$$ADE = \sum_{i=1}^{N} \sum_{t=t_{sta}}^{t_{fin}} \frac{\sqrt{(P_x^{i,t} - G_x^{i,t})^2 + (P_y^{i,t} - G_y^{i,t})^2}}{N(t_{fin} - t_{sta})}, \tag{8}$$

where $P_x^{i,t}$ and $P_y^{i,t}$ denote the predicted movement along the x-axis and y-axis, respectively, of the ith pedestrian during the next 0.5 s from time t, and $G_x^{i,t}$ and $G_y^{i,t}$ are the corresponding ground truths. The mean ADE of all the previous repetitions is calculated and recorded.

With more than 1999 repetitions, the variation of ADE with the increase of hidden neurons will reach a stable trend, and a U-shape characteristic can be observed, so if the number of repetitions exceeds 1999, the repetition will terminate, and the mean ADE of all the repetitions for n neurons in the hidden layer will be recorded as A_n. The effectiveness is tested with the number of hidden neurons ranging from 1 to 70, and the number of neurons at which the movement is predicted with the best accuracy is selected as the number of neurons in the hidden layer.

Parameter analysis

After determining the input variables, output variables, and number of neurons in the hidden layer, the actual structure of the neural network model is determined. Garson's method is capable of quantifying the relative importance and select variables (Papatheocharous and Andreou 2010; Fischer 2015; Yousefi et al. 2018; Liu et al. 2018). Other methods such as Olden's method and SHapley Additive exPlanations (SHAP) values can quantify the intensity and direction of each input variable's contribution to each output variable (Olden and Jackson 2002; Lundberg and Lee 2017), which is helpful to understand the relationship between each input and output variable in the model. Since both Olden's method and Garson's method are weight-based methods, if Olden's method is used, comparing the relative importance and contribution can show the role of specific variables in the structure of neural network. For a neural network model with N neurons in the input layer and L neurons in the hidden layer, the contribution of the ith input variable to the kth output variable (C_{ik}) can be defined as

$$C_{ik} = \sum_{j=1}^{L} \omega_{ij} v_{jk}, \tag{9}$$

where ω_{ij} is the weight of the ith neuron in the input layer and jth neuron in the hidden layer, and v_{jk} is the weight of the jth neuron in the hidden layer and kth neuron in the output layer. C_x and C_y are the contributions of each input variable to N_x and N_y, respectively. The magnitude indicates the strength of contribution and the sign shows the direction of contribution. The group of input variables will not be further changed by the result of C_x and C_y in this step, so no threshold is specified for C_x and C_y. In this study, the five variables that have the largest magnitude of contribution to N_x and N_y are selected as the variables that contribute significantly to the prediction of N_x and N_y respectively. Olden's method can also reach a stable trend with 999 repetitions (Olden and Jackson 2002), so the final neural network runs 999 repetitions using random initial weights. To evaluate the performance of the final model, the average of R-square, ADE, relative importance and contribution will be calculated.

Case study

Description of the site

The area chosen for case study is a busy commercial district near the Flood Control Victory monument at the Harbin Central Street, China, as shown in Fig. 9. This district is a pedestrian-only space, and the area of interest in this study is 59 m long and 9 m wide.

Fig. 9 Representation of a scene from the case-study area

This region contains six dominant starting points and destinations: (1) left side, (2) slopes and steps below, (3) right side, (4) entrance of a shopping mall, and (5 + 6) two restaurant entrances. Except for the left and right sides, this area is surrounded by building facades, handrails, steps, and slopes, which form the borders of this area.

Using an unmanned aerial vehicle (UAV), a video of length 10 min was recorded on Saturday, September 1, 2018, which was a pleasant sunny day to visit this area. The camera captured the video at a resolution of 2720×1536 pixels and frame rate of 24 fps, and was attached to a DJI Mavic Pro light UAV weighing 743 g. The flying height of the UAV was 90 m, which was sufficiently high to observe both ends of the area of interest and sufficiently low to be able to clearly recognise pedestrian movements.

In the 10 min video, 1565 pedestrians appeared in the region of interest. In this study, pedestrians who are obviously smaller than other pedestrians are regarded as children and pedestrians on crutches and wheelchairs are regarded as needing other people's assistance in walking. Consequently, most people could walk and choose their route independently with only a few children and pedestrians who need other people's assistance in walking. During the video-recording period, all the pedestrians walked calmly without exhibiting any abnormal behaviour.

Identification of social groups

As the present study considers the walking behaviour of members in social groups, pedestrians having an abnormal average walking speed (less than 0.5 m/s) are ignored. The grouping process can be divided into two stages. By means of a set of heuristics, pedestrians that have similar trajectories are identified in the first stage. Then, in the second stage, the results of the first stage are reviewed manually, and the social groups are identified.

The first stage consisted of six steps performed automatically by a program, as shown in Fig. 10; five of these steps relate to the criteria used to determine the similarity between pedestrians, and the remaining step relates to identifying possible groups based on the result of the similarity analysis. In the first five steps, the trajectory of every pedestrian is compared with that of all the other pedestrians who appear in the camera image simultaneously. Any pair that meets all the criteria is regarded as being part of the same group. First, among the six predetermined starting points in the area of interest, as shown in Fig. 9, pedestrians sharing the same starting point as the specific pedestrian being analysed are selected and moved to the next step. Similarly, pedestrians having the same destination are identified in the second step. Accordingly, the distance between these pairs of pedestrians

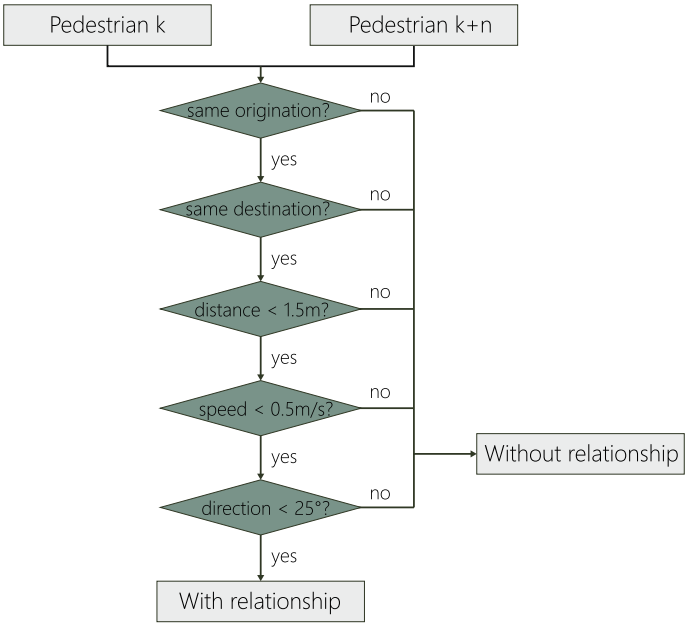


Fig. 10 Automatic identification of social groups by a program

in the entire trajectory is calculated. The criterion for relationship is met if the mean value of the distance is less than 1.5 m (Li et al. 2015). The fourth step considers the difference in speed; a pair is retained if the mean value of the speed difference between these two pedestrians is less than 0.5 m/s (Li et al. 2015). The last step of determining the relationship is based on the difference in the walking direction; a pair is considered to be related if the mean value is less than 25° (Li et al. 2015). The sixth step is to identify all the members of a group. In this step, all pairs with a relationship are first regarded as a two-person group. If *Pedestrian a* in *Group A* is related to *Pedestrian b* in *Group B*, *Group A* and *Group B* are combined, and both the original groups are deleted. The initial groups are combined step-by-step until none of these groups can be combined further.

In the second stage, members excluded from the social groups in the previous step are added to the group, and irregular groups are disregarded. There is one criterion for adding pedestrians to a group and three criteria for neglecting a group. The pedestrians or groups that have apparent interactions with anyone in a predetermined group (e.g. touching, waiting, turning body to talk) in the video are added to that group. In this study, three types of behaviours are treated as irregular. First, certain factors such as a food stand or dustbin attract specific pedestrians; it is difficult to incorporate this in a model that considers all pedestrians. Therefore, groups attracted by these factors are disregarded. Second, the groups with children are disregarded because children tend to behave more casually than adults, and their movements influence the movement of other individuals in the group. Third, the groups stopping to find their way or walking back after reaching the intended destination are disregarded. In such conditions, the deceleration or return is not induced by the factors considered in this research. The grouping result is achieved by considering these manual adjustments. In this stage, 58 groups added new members because of interactions observed, and 48 irregular groups were disregarded.

Trajectory extraction and grouping

The UAV captured the video featuring images of 2720 × 1536 pixels. Because the pan-tilt camera can prevent video shaking, the averaged measurement error in every time interval (0.5 s) is diminished to 0.0173 m. The video analysis software 'Tracker', which is a program built on an open-source Physics Java framework, was used to derive the trajectories of all pedestrians within the study area from the video recordings. The origin of the coordinate system, the positive direction of the x-axis, and the unit length can be fixed on the unchanged background in the video (the building facade in this study), which can further reduce the averaged measurement error. Combined with the unit length in the video, this study applies perspective transformation to convert the data to GPS level. Owing to the high altitude of the UAV and the low number of pixels per individual, the movements of the upper body were part of the captured trajectories. After performing the semi-automatic tracking procedure, the tracking errors produced by the software were corrected manually by visual inspection.

Most pedestrians were part of various social groups in the video; the grouping condition is shown in Fig. 11. There were 805 pedestrians walking normally in social groups, which incorporated 34,005 data points that can be used for training and testing. The largest group observed in the video consisted of 6 people, and most of the pedestrians belonged to groups with 2 or 3 people, accounting for 50% and 30% of all pedestrians, respectively; this result corresponds to the result of James (1953). The remaining groups consisted of pedestrians walking individually, pedestrians moving slowly (at less than 0.5 m/s), and members in the groups demonstrating behaviour triggered by factors that are not considered in this study, such as those attracted by other factors or those who had lost their way. Members in social groups ought to walk together; however, some larger groups split into sub-groups when the pedestrian density nearby was high and reunited after passing the crowded area. Although social groups may split, their members are more likely to walk together within a certain range of distance.

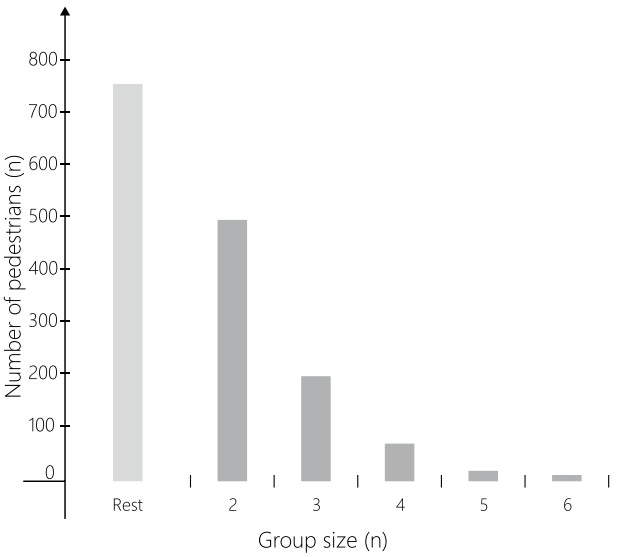


Fig. 11 Number of groups in different group sizes

Results

In the first procedure, the R-square of 18 variables are larger than 0.003 and considered as having strong prediction ability in this study. After the second procedure, 16 variables are found to influence the acceleration of pedestrians that are part of social groups when all variables having strong prediction ability are used as input variables. The relative importance (I) and contribution (C_x and C_y) of these 16 variables in the final model are shown in Table 1. These variables feature the current movement characteristics, pedestrian–environment interactions, inter-personal interactions, intra-group relationships, and inter-group variables. The influence of the variables included in each group is discussed per group.

Using these 16 variables as inputs, the mean ADE of the training and testing process of 1–70 neurons in the hidden layer is shown in Fig. 12. In general, the training accuracy increases with increase in the number of neurons in the hidden layer. Furthermore, the prediction accuracy increases with increase in the number of neurons in the hidden layer when the number is not large; however, after a threshold, the ADE increases because of overfitting. Therefore, the neural network featured 17 neurons in the hidden layer, which can exhibit the best prediction accuracy.

Movement characteristics

This group of variables consists of four variables, namely, current velocity (V), angle between the current movement direction and target direction (D), acceleration along the x-axis (A_x), and acceleration along the y-axis (A_y). Table 2 shows the result of the first variable selection procedure, which is the R-square between the prediction result of each variable and the ground truth of acceleration along the x-axis and y-axis. The prediction results of V, D, A_x, and A_y are strongly correlated with acceleration along the x-axis (N_x).

Table 1 Relative importance of all input variables, where I represents relative importance, C_x represents the contribution to N_x, and C_y represents the contribution to N_y

Variables	I	C_x	C_y
V	0.1377	− 1.0525	0.0044
D	0.2023	0.2487	0.8056
A_x	0.2724	0.3180	− 0.0273
A_y	0.3197	− 0.0280	1.6436
$B_d^{0,30}$	0.1017	0.1928	− 0.0204
$B_d^{30,60}$	0.1352	0.1503	0.2502
$B_d^{60,90}$	0.1581	− 0.2976	− 0.0554
$B_d^{90,120}$	0.1162	0.0839	− 0.1110
$T_s^{0,30}$	0.0656	− 0.0453	− 0.1510
$T_d^{0,30}$	0.0742	− 0.0243	− 0.4749
$T_d^{30,60}$	0.0398	0.0261	− 0.1696
G_o	0.0657	0.1716	0.1793
G_a	0.0434	− 0.0273	0.3694
G_x	0.0875	0.1718	− 0.0439
G_y	0.1204	0.0326	− 0.0179
$L_s^{0,15}$	0.0601	0.0610	− 0.0728

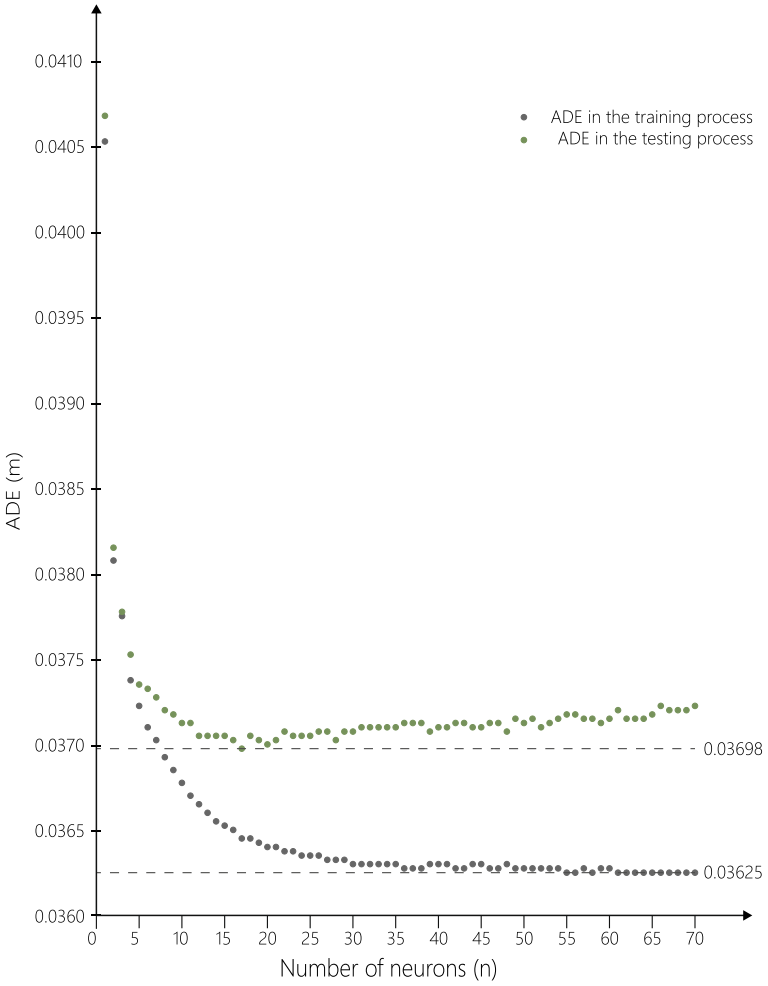


Fig. 12 Average displacement error in the training and testing process with 1–70 neurons in the hidden layer

Table 2 R-square between prediction results using a single variable describing movement characteristics as input and the ground truth of the testing data

Variables	*Rx*	*Ry*
V	0.0084	0.0000
D	0.0031	0.1288
A_x	0.1867	− 0.0007
A_y	0.0115	0.2015

Moreover, the prediction results of D and A_y are correlated with the acceleration along the y-axis (N_y).

In the second variable selection procedure, all the four variables are incorporated in the final model, as shown in Table 1. The fact that V, A_x, and A_y are incorporated in the

prediction model shows that the walking condition in the past 0.5 s strongly influences the movement during the next 0.5 s. The negative contribution from V to N_x indicates that pedestrians attempt to maintain their walking speed within a desirable range. Both the contributions of A_x to N_x and A_y to N_y are positive, indicating that pedestrians follow a similar turning trend as adopted in the previous time interval. The contribution of D to N_y is positive, indicating that pedestrians try to walk in a similar direction to the destination. N_y is not correlated with V, which could be because fast-walking pedestrians make smaller directional changes (Duives et al. 2017); however, a small change in the walking direction induces a large N_y. The correlation between acceleration and V and D is consistent with previous research results (Helbing and Molnár 1995; Zanlungo et al. 2014), and this study validates a continuous trend of acceleration.

Pedestrian–environment interaction

Of the 12 variables describing the pedestrian–environment interaction, $B_d^{0,30}$, $B_d^{30,60}$, $B_d^{60,90}$, and $B_d^{90,120}$ are found to be correlated with N_y in the first procedure, as shown in Table 3. This finding implies that obstacles in the visible area have a strong influence on the variation of walking dynamics.

In the final model, all the four variables are incorporated, as shown in Table 1. $B_d^{30,60}$, $B_d^{60,90}$, and $B_d^{90,120}$ contribute more to N_y than $B_d^{0,30}$, which can be caused by the infrequent occurrence of walls on pedestrian routes. The contribution of $B_d^{30,60}$ is positive, which can be the result of the guiding effect of the wall. A previous study found that pedestrians tend to walk along walls to get a smooth path (Lee 2015). The negative effects of $B_d^{60,90}$ and $B_d^{90,120}$ ensure a safe distance between pedestrians and walls. If the specific pedestrian no longer walks along the wall, resulting in the wall exceeding 30°–60°, the repulsive force of the wall will be the dominant factor. This finding validates the guiding effect found by Lee (2015) and the repulsive force found by Johansson et al. (2007), which are combined in the neural network model.

Inter-personal interaction

The neural network model includes the time to an upcoming inter-pedestrian collision and the distance to surrounding pedestrians. In the first procedure, $T_s^{0,30}$ and $T_d^{0,30}$ are found to be correlated with N_x, and $T_d^{0,30}$, $T_d^{30,60}$, $D_d^{90,120}$ and $D_d^{120,150}$ are correlated with N_y, as shown in Table 4.

$T_s^{0,30}$, $T_d^{0,30}$, and $T_d^{30,60}$ are incorporated in the prediction model, as shown in Table 1. In the prediction model, the C_x of $T_s^{0,30}$ is negative. This correlation shows that face-to-face collisions decrease the walking speed. The influence of $T_d^{0,30}$ on N_x is also negative but weaker than that of $T_s^{0,30}$—this indicates that the deceleration will be more apparent for

Table 3 R-square between prediction results using a single variable describing pedestrian–environment interaction as input and the ground truth of the testing data

Variables	*Rx*	*Ry*
$B_d^{0,30}$	− 0.0004	0.0169
$B_d^{30,60}$	− 0.0005	0.0116
$B_d^{60,90}$	− 0.0004	0.0069
$B_d^{90,120}$	− 0.0004	0.0044

Table 4 R-square between prediction results using a single variable describing inter-personal interaction as input and the ground truth of the testing data

Variables	*Rx*	*Ry*
$T_s^{0,30}$	0.0097	− 0.0001
$T_d^{0,30}$	0.0067	0.0085
$T_d^{30,60}$	− 0.0003	0.0059
$D_d^{90,120}$	− 0.0005	0.0124
$D_d^{120,150}$	− 0.0006	0.0069

collisions approaching from the left. This result may be caused by the fact that most pedestrians walk on the right side, resulting in less space to evade collisions from the left. For N_y, the C_y of both $T_d^{0,30}$ and $T_d^{30,60}$ is negative, and the C_y of $T_d^{0,30}$ is stronger. These correlations show that in face-to-face collisions, pedestrians turn in the opposite direction to evade the collision, and the collision nearer to the current walking path causes a larger impact. It can also be found that in the final model, the C_y of $T_s^{0,30}$ is negative. This can be attributed to the phenomenon that in face-to-face collision, pedestrians are more prone to turn right to avoid, which is consistent with the previously found right-hand behaviour (Moussaïd et al. 2009). The relative importance of the time to collision and the distance of collision is compared in this study. The exclusion of $D_d^{90,120}$ and $D_d^{120,150}$ shows that pedestrians are more sensitive to the time to collision than the distance to surrounding pedestrians.

Intra-group relationship

Five variables in the neural network model describe intra-group relationships, namely, group size (G_s), longitudinal distance (G_o), lateral distance (G_a), difference in speed along the x-axis (G_x), and difference in speed along the y-axis (G_y). G_o, G_a, G_x, and G_y are selected in the first procedure, and the result is presented in Table 5.

All four variables are incorporated in the final model, as shown in Table 1. The C_x of G_o and G_x is stronger, and the C_y of G_o and G_a are stronger. The C_x of G_o and C_y of G_a are positive. These two correlations indicate cohesion in social groups, which causes pedestrians to maintain proximity with their group members. The C_x of G_x is also positive—this indicates velocity coordination, wherein pedestrians in the same social group attempt to maintain a similar walking speed as their fellow group members. The contribution from G_y is slightly different from the finding of velocity coordination proposed by Qiu and Hu (2010). The less obvious contribution and large relative importance of G_y can be caused by the complex influence of G_y in this study. In some cases, the affected pedestrians react before their group members; Under other influences, their team members react first. Although the influence of G_y is important, it depends on the conditions described by other variables.

Table 5 R-square between prediction results using a single variable describing intra-group relationship as input and the ground truth of the testing data

Variables	*Rx*	*Ry*
G_o	0.0086	− 0.0003
G_a	− 0.0008	0.0113
G_x	0.0283	− 0.0009
G_y	− 0.0003	0.0443

Table 6 R-square between prediction results using a single variable describing inter-group relationship as input and the ground truth of the testing data

Variables	*Rx*	*Ry*
$L_s^{0,15}$	0.0036	− 0.0012

Table 7 Comparison of ADE and FDE of other methods with that of our neural network model

Models	*ADE*	*FDE*
Discrete model	1.0701	2.0132
SF model	1.4717	2.2984
Social LSTM	0.6326	1.3255
Our neural network	0.3559	0.6863

Inter-group relationship

Six variables describe the leader–following behaviour, and the prediction result of $L_s^{0,15}$ is found to be correlated with N_x, as shown in Table 6. $L_s^{0,15}$ is incorporated in the final model. The C_x of $L_s^{0,15}$ is positive, indicating that pedestrians accelerate and decelerate according to the leader, and the leaders near the current walking path have a significant influence. The effect of this influencing factor is similar to that of Robin et al. (2009), and this study further defines the dominant scope of pedestrian impact.

Neural network model

The final neural network model contains 16 input variables, namely, V, D, A_x, A_y, $B_d^{0,30}$, $B_d^{30,60}$, $B_d^{60,90}$, $B_d^{90,120}$, $T_s^{0,30}$, $T_d^{0,30}$, $T_d^{30,60}$, G_o, G_a, G_x, G_y, and $L_s^{0,15}$. To further validate the effectiveness of our neural network model, the prediction performance with the same datasets of state-of-the-art models, namely, a discrete model (Robin et al. 2009), the SF model (Zanlungo et al. 2014), and the Social LSTM model (Alahi et al. 2016), is compared with that of our neural network model. The dataset is separated into the first five minutes and last five minutes. For the LSTM model and our neural network model, the first half is used for training, and the second half is used for testing. The discrete model and SF model predict the second half. Because Social LSTM model predicts for 12 steps, in order to have an objective comparison between methods, every model predicts for 12 steps.

The average displacement error (ADE) and final displacement error (FDE) in the prediction of 12 steps (6 s) are listed in Table 7. FDE is calculated as the distance between the predicted final destination and the true final destination at end of the prediction period (Alahi et al. 2016). The prediction accuracy decreases with the increase of prediction steps, so the ADE of predicting 12 steps is significantly larger than that of predicting 1 step.

The results conform to the finding that data-based methods show better prediction accuracy (Kothari et al. 2020). Our proposed neural network model achieves a better prediction accuracy than Social LSTM, which may be a result of our comprehensive set of input variables. Our model has four main advantages. First, the long-term goal of pedestrians is not included in Social LSTM. In our proposed neural network, the variable D describes

the direction of destination of a specific pedestrian. Second, studies on human trajectory prediction can be categorized into the learning of human–human (social) interactions, human–space (physical) interactions, or both (Kothari et al. 2020). Social LSTM learns human–human (social) interactions, and our proposed model learns both human–human (social) interactions and human–space (physical) interactions. Third, as a non-grid prediction model, our proposed neural network model can incorporate distant influences on a specific pedestrian. Fourth, in our proposed model, the social groups are identified by both a program and observation, which can further improve the prediction accuracy.

The relative importance and contribution in the final model are presented in Table 1. In the neural network model, the five most important input variables are A_y, A_x, D, $B_d^{60,90}$ and V. This shows that the current walking condition has the greatest impact on the walking dynamics. V, A_x, $B_d^{60,90}$, D, and $B_d^{0,30}$ are the most influential factors in the prediction of N_x, and A_y, D, $T_d^{0,30}$, G_a, and $B_d^{30,60}$ have the largest contribution to N_y.

The ADE of the final model to predict the next time step is slightly lower than 0.0370, and the R-square of predicting the variation of speed in x-axis and y-axis are 0.3161 and 0.4048 respectively. Furthermore, the input variable groups are excluded from the neural network in individual succession to calculate the ADE of prediction result with the remaining variables, as shown in Table 8. When the movement characteristics group is excluded from the input variables, the ADE of the prediction result is largest; the ADE is second-largest when the intra-group relationship group is excluded. Previous studies focused on building models with better prediction accuracy, and the relative importance of different variables in a comprehensive model is not directly compared. The results of this study are consistent with previous studies, that is, the model incorporating group behaviour can improve the prediction accuracy (Moussaïd et al. 2010; Bisagno et al. 2019). Furthermore, by comparing the importance of group behaviour with other variables, this study verifies the importance of group behaviour in the comprehensive model, and found that movement characteristics and intra-group relationship are the most effective of all input variables.

Conclusion

Although attempts have been made by some researchers to incorporate group dynamics into microscopic pedestrian models, the principle of interaction within social groups remains unclear. This study proposed a neural network model that can clarify the relative importance and contribution of each influential factor on the prediction of walking dynamics of pedestrians in social groups.

This study presented a three-step method to determine a neural network model for pedestrian walking dynamics. The first two steps aim to select the input variables, and the third step determines the number of neurons in the hidden layer. Because the movement of

Table 8 ADE of our neural network, obtained by excluding input variable groups in individual succession

Excluded group	*ADE*
Movement condition	0.0442
Pedestrian–environment	0.0370
Inter-pedestrian	0.0373
Intra-group	0.0382
Inter-group	0.0371

pedestrians is two-dimensional (x-axis and y-axis in the coordinate system), an input variable that is correlated with the prediction of the walking dynamics in any axis is selected in the first step. In the second step, all the variables selected in the first step are input variables, and those determined as insensitive in the prediction model are sequentially excluded using Garson's method, until all the variables are sufficiently sensitive to be incorporated in the final model. In the third step, the ADE of the prediction model with an increasing number of neurons in the hidden layer is calculated, and the number of neurons yielding the best prediction accuracy is selected.

It is found that the movement dynamics of pedestrians in social groups is influenced by the group members. Both cohesion and velocity coordination in social groups are influential factors making non-negligible contributions to the variation in the current velocity. In this study, the factor with the strongest influence on the walking dynamics was velocity coordination, which is related to group behaviour. Pedestrians tend to maintain a similar speed as their group members along the x-axis, but the velocity coordination along the y-axis is complex and depends on the specific conditions. The cohesion in social groups influences the members of the group, whereby they strive to remain close to the centre of their group. Consequently, pedestrians change their velocity to maintain proximity to the centre. In this study, the influence of cohesion and velocity coordination was validated, and the weight of the influence of group dynamics was quantified.

The same type of parameter can exhibit different types of influence on the variation in walking dynamics, and some influences can even be contrary, as observed in the pedestrian–environment interaction in this study. Obstacles are generally considered a type of repulsion, and the strength of repulsion corresponds to the distance and angle with respect to the pedestrian. In this study, it was found that the force from obstacles within the range of 60°–120° diverts the pedestrians to the opposite direction; however, in the case of obstacles within the range of 30°–60°, the boundary of the area attracts the pedestrians. Consequently, some correlations cannot be simply regarded as attraction or repulsion, because the influence can be more complex.

For pedestrians in social groups, the time to collision shows a greater contribution than the distance to surrounding pedestrians, which has been extensively considered in previous models. This group of variables incorporates all types of collisions, including colliding pedestrians in the same and opposite directions. Therefore, the influences of both surpassing behaviour and avoidance behaviour can be incorporated to find the correlation with variation in the original walking dynamics. The strength and category of influence are determined by the time remaining for an upcoming collision and the direction from which the collider approaches.

Applications and future work

This study investigated a comprehensive neural network model to predict pedestrian movements in social groups. This study is important in the design of urban spaces, especially spaces intended for commercial or entertainment activities. Previous models mainly considered the parameters influencing individual behaviour. However, pedestrians in these spaces are more liable to be in social groups, where their behaviour is very different from those of individual pedestrians. Because the constitution of different sizes of social groups is related to the activities in which the pedestrians participate, the positions of bottleneck areas, crowd distribution, and pedestrian capacity are influenced by the function of the

space. Being aware of the properties of the space to be designed, rather than simply imagining the use of the space, can enable designers to incorporate the environmental settings and group composition of that space into our neural network model to predict the use of the space and identify bottlenecks or void areas before completing the project. Furthermore, changes in the ownership or characteristics of a specific area cause a corresponding change in the amount and percentage of social groups in the area. This model can also be used to evaluate the influence of renovation design on pedestrian walking dynamics and ensure that the renovated space is suitable for its new function.

In this study, the importance of including group behaviour was examined. Group dynamics has been studied by several researchers, all of whom have proposed the idea that group dynamics is important in crowd dynamics. However, its relative importance in comparison with other types of influences has not been quantified. In this study, it was found that the inclusion of intra-group relationship is highly important, except in the current walking condition. Future studies that aim to optimise existing models such as the SF model, Markov model, and LSTM model should consider the relative speed and position of a specific pedestrian moving in a social group to predict the movement dynamics of pedestrians in real life, and their influences should be incorporated.

Group behaviour and many other factors could potentially influence the walking dynamics of pedestrians. Therefore, a variety of input variables can be added synchronously in a comprehensive prediction model, which causes the model to contain useless information and increases the model complexity significantly, especially in the case of models such as the Markov or LSTM model. Selecting a part of the variables before determining the final group of input variables is important in the first step. It is also meaningful to test the relative importance of a specific parameter before adding it to the SF model. Many existing models require assumptions to be made by the modeller (Cheng et al. 2014); however, determining whether these parameters are influential and crucial is difficult. This study provides a method to select variables that are most influential in walking dynamics prediction models. Therefore, this study makes a fundamental contribution to the selection of influencing factors and the quantification of the influence to simplify the input variables and identify non-negligible influences.

In the future, the model can be further improved in the following aspects. First, the input variables can be made flexible; that is, if additional influential factors are found to be important, the model should allow these factors to be added into the input variable group to test if they can further improve the accuracy. Second, other methods could also be effective in the variable selection process, such as SHAP values. If new methods are sufficiently verified to be effective in variable selection of other machine learning models, such as neural network with multiple hidden layers and LSTM model, different variable selection processes and models can be compared to further improve the prediction accuracy. Third, because different kinds of interaction in social groups caused by the group members' gender, age, and relationship can affect the influence of social groups, the group dynamics in a specific kind of space can have some unique features. In addition, other factors such as surrounding density and time of day can also influence group dynamics, and the focus of this study is the calm walking mode. The multinomial logit model can be further improved by training the neural network based on experiments in other spaces and conditions. Fourth, further research can use our neural network model to predict the pedestrian distribution in typical urban environment, and provide planners with more generic and concrete design strategies, such as in the form of tables. Fifth, the video used in this study was shot from the top, which is suitable for tracking pedestrians; to further improve the prediction accuracy, another camera can be used to capture and add other information to the model. On

one hand, the step frequency is unobservable, making the variation caused by pedestrian locomotion modes with steps unquantifiable. The effects of the swing of the body can be eliminated by adding a video that can observe the step frequency. On the other hand, the trajectories of pedestrians are at least 9 m, and the average density is only 0.1938 p/m^2, making it feasible to identify social groups because, under the influence of the territory effect, most pedestrians do not stay and walk together throughout the trajectory if they are not from the same social group (Helbing and Molnár 1995). However, it still cannot be ensured that all social groups are identified without errors. Thus, future attempts should add another video to record visual hints such as talking and appearance to further improve the accuracy of identifying social groups.

Acknowledgements The authors would like to thank Ying Liu, Yan Feng, Junfeng Fu, and Yongjie Pan for their assistance in conducting this research.

Authors' contribution SS: conceptualization, data curation, formal analysis, methodology, software, roles/writing—original draft. CS: conceptualization, data curation, supervision, writing—review and editing. DCD: conceptualization, formal analysis, methodology, software, roles/writing—original draft. SPH: conceptualization, supervision, writing—review and editing.

Funding This work was supported by the China Scholarship Council [Grant Number 201806120266]; the research program 'Allegro: Unravelling slow mode travelling and traffic: with innovative data to a new transportation and traffic theory for pedestrians and bicycles' [ERC Grant Agreement No. 669792]; and the National Natural Science Foundation of China [Grant Number 51878202].

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Li, F., Savarese, S.: Social LSTM: Human trajectory prediction in crowded spaces. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 961–971 (2016). https://doi.org/10.1109/CVPR.2016.110

Ashbrook, D., Starner, T.: Using GPS to learn significant locations and predict movement across multiple users. Pers. Ubiquitous Comput. **7**, 275–286 (2003). https://doi.org/10.1007/s00779-003-0240-0

Bakeman, R., Beck, S.: The size of informal groups in public. Environ. Behav. **6**, 378–390 (1974). https://doi.org/10.1177/001391657400600305

Barten, A.P.: The Coefficient of Determination for Regression Without a Constant Term, pp. 181–189. Springer, Dordrecht (1987). https://doi.org/10.1007/978-94-009-3591-4_12

Bartoli, F., Lisanti, G., Ballan, L., Bimbo, A.: Context-aware trajectory prediction. In: 2018 24th International Conference on Pattern Recognition (ICPR), pp. 1941–1946 (2018). https://doi.org/10.1109/ICPR.2018.8545447

Bisagno, N., Zhang, B., Conci, N.: Group LSTM: group trajectory prediction in crowded scenarios. In: Computer Vision—ECCV 2018 Workshops, vol 11131, pp. 213–225 (2019). https://doi.org/10.1007/978-3-030-11015-4_18

Burkert, F., Bamler, R.: Graph-based analysis of pedestrian interactions and events using hidden Markov models. Photogrammetrie - Fernerkundung - Geoinf. **2012**, 701–710 (2012). https://doi.org/10.1127/1432-8364/2012/0150

Chella, A., Frixione, M., Gaglio, S.: Understanding dynamic scenes. Artif. Intell. **123**, 89–132 (2000). https://doi.org/10.1016/S0004-3702(00)00048-5

Cheng, L., Yarlagadda, R., Fookes, C., Yarlagadda, P.K.D.V.: A review of pedestrian group dynamics and methodologies in modelling pedestrian group behaviours. World J. Mech. Eng. **1**, 1–13 (2014)

Duives, D.C., Daamen, W., Hoogendoorn, S.P.: State-of-the-art crowd motion simulation models. Transp. Res. C Emerg. Technol. **37**, 193–209 (2013). https://doi.org/10.1016/j.trc.2013.02.005

Duives, D.C., Daamen, W., Hoogendoorn, S.P.: Influence of group size and group composition on the adhered distance headway. Transp. Res. Procedia **2**, 183–188 (2014). https://doi.org/10.1016/j.trpro.2014.09.026

Duives, D.C., Daamen, W., Hoogendoorn, S.P.: Operational walking dynamics of crowds modeled with linear regression. Transp. Res. Rec. **2623**, 90–97 (2017). https://doi.org/10.3141/2623-10

Elmchichi, L., Belhassan, A., Aouidate, A., Ghaleb, A., Lakhlifi, T., Bouachrine, M.: QSAR study of new compounds based on 1,2,4-triazole as potential anticancer agents. Phys. Chem. Res. **8**(1), 125–137 (2020). https://doi.org/10.22036/PCR.2019.204753.1685

Fajen, B.R., Warren, W.H.: Behavioral dynamics of steering, obstacle avoidance, and route selection. J. Exp. Psychol. Hum. Percept. Perform. **29**, 343–362 (2003). https://doi.org/10.1037/0096-1523.29.2.343

Feng, Y., Li, D.: An empirical study and a conceptual model on heterogeneity of pedestrian social groups for friend-group and family-group. In: Proceedings of Pedestrian and Evacuation Dynamics 2016, pp. 402–407 (2016)

Fernando, T., Denman, S., Sridharan, S., Fookes, C.: Soft + hardwired attention: an LSTM framework for human trajectory prediction and abnormal event detection. Neural Netw. **108**, 466–478 (2018). https://doi.org/10.1016/j.neunet.2018.09.002

Fernando, T., Denman, S., Sridharan, S., Fookes, C.: GD-GAN: generative adversarial networks for trajectory prediction and group detection in crowds. In: Computer Vision—ACCV 2018, vol. 11136, pp. 314–330 (2019). https://doi.org/10.1007/978-3-030-20887-5_20

Festa, A., Tosin, A., Wolfram, M.T.: Kinetic description of collision avoidance in pedestrian crowds by sidestepping. Kinet. Relat. Models **11**, 491 (2018). https://doi.org/10.3934/krm.2018022

Fischer, A.: How to determine the unique contributions of input-variables to the nonlinear regression function of a multilayer perceptron. Ecol. Model. **309–310**, 60–63 (2015). https://doi.org/10.1016/j.ecolmodel.2015.04.015

Fu, Y., Li, M., Liang, J., Hu, X.: Modeling and simulating the walking behavior of small pedestrian groups. AsiaSim **474**, 1–14 (2014). https://doi.org/10.1007/978-3-662-45289-9_1

Gambs, S., Killijian, M.O., del Prado Cortez, M.N.: Show me how you move and I will tell you who you are. In: Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS, vol. 4, pp. 34–41 (2010). https://doi.org/10.1145/1868470.1868479

Gambs, S., Killijian, M.O., del Prado Cortez, M.N.: Next place prediction using mobility Markov chains. In: Proceedings of the First Workshop on Measurement, Privacy, and Mobility, pp. 1–6 (2012). https://doi.org/10.1145/2181196.2181199

Garson, G.D.: Interpreting neural-network connection weights. AI Expert **6**, 46–51 (1991)

Ge, W., Collins, R.T., Ruback, R.B.: Vision-based analysis of small groups in pedestrian crowds. IEEE Trans. Pattern Anal. Mach. Intell. **34**, 1003–1016 (2012). https://doi.org/10.1109/TPAMI.2011.176

Heaton, J., McElwee, S., Fraley, J., Cannady, J.: Early stabilizing feature importance for tensorflow deep neural networks. In: 2017 International Joint Conference on Neural Networks (IJCNN), pp. 4618–4624 (2017). https://doi.org/10.1109/IJCNN.2017.7966442

Helbing, D., Molnár, P.: Social force model for pedestrian dynamics. Phys. Rev. E **51**, 4282–4286 (1995). https://doi.org/10.1103/PhysRevE.51.4282

Helbing, D., Molnár, P., Farkas, I.J., Bolay, K.: Self-organizing pedestrian movement. Environ. Plan. B. Plan. Des. **28**, 361–383 (2001). https://doi.org/10.1068/b2697

James, J.: The distribution of free-forming small group size. Am. Sociol. Rev. **18**, 569–570 (1953). https://doi.org/10.2307/2087444

Johansson, A., Helbing, D., Shukla, P.K.: Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. Adv. Complex Syst. **10**, 271–288 (2007). https://doi.org/10.1142/S0219525907001355

Kiefer, A., Rio, K., Bonneaud, S., Walton, A., Warren, W.: Quantifying and modeling coordination and coherence in pedestrian groups. Front. Psychol. **8**, 1–13 (2017). https://doi.org/10.3389/fpsyg.2017.00949

Koh, P.P., Wong, Y.D.: Influence of infrastructural compatibility factors on walking and cycling route choices. J. Environ. Psychol. **36**, 202–213 (2013). https://doi.org/10.1016/j.jenvp.2013.08.001

Köster, G., Seitz, M., Treml, F., Hartmann, D., Klein, W.: On modelling the influence of group formations in a crowd. Contemp. Soc. Sci. **6**, 397–414 (2011). https://doi.org/10.1080/21582041.2011.619867

Kothari, P., Kreiss, S., Alahi, A.: Human trajectory forecasting in crowds: a deep learning perspective, pp. 1–33. arXiv:2007.03639 (2020)

Lakes, T., Müller, D., Krüger, C.: Cropland change in southern Romania: a comparison of logistic regressions and artificial neural networks. Landsc. Ecol. **24**, 1195–1206 (2009). https://doi.org/10.1007/s10980-009-9404-2

Lee, S.J.: Navigational pedestrian movement model with vision-driven agents. J. Asian Archit. Build. Eng. **14**, 371–378 (2015). https://doi.org/10.3130/jaabe.14.371

Li, X., Duan, P., Zheng, S., Li, B., Liu, M.: A study on the dynamic spatial-temporal trajectory features of pedestrian small group. In: 2015 2nd International Symposium on Dependable Computing and Internet of Things (DCIT), pp. 112–116 (2015). https://doi.org/10.1109/DCIT.2015.9

Liu, J., Boyle, L.N., Banerjee, A.G.: Predicting interstate motor carrier crash rate level using classification models. Accid. Anal. Prev. **120**, 211–218 (2018). https://doi.org/10.1016/j.aap.2018.06.005

Lu, L., Ren, G., Wang, W., Wang, Y.: Modeling walking behavior of pedestrian groups with floor field cellular automaton approach. Chin. Phys. B **23**, 088901 (2014). https://doi.org/10.1088/1674-1056/23/8/088901

Lu, L., Chan, C., Wang, J., Wang, W.: A study of pedestrian group behaviors in crowd evacuation based on an extended floor field cellular automaton model. Transp. Res. C Emerg. Technol. **81**, 317–329 (2017). https://doi.org/10.1016/j.trc.2016.08.018

Lundberg, S.M., Lee, S.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, vol. 30, pp. 4765–4774 (2017)

Moussaïd, M., Helbing, D., Garnier, S., Johansson, A., Combe, M., Theraulaz, G.: Experimental study of the behavioural mechanisms underlying self-organization in human crowds. Proc. R. Soc. B **276**(1668), 2755–2762 (2009). https://doi.org/10.1098/rspb.2009.0405

Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., Theraulaz, G.: The walking behaviour of pedestrian social groups and its impact on crowd dynamics. PLoS ONE **5**, e10047 (2010). https://doi.org/10.1371/journal.pone.0010047

Nascimento, J.C., Figueiredo, M.A.T., Marques, J.S.: Trajectory classification using switched dynamical hidden Markov models. IEEE Trans. Image Process. **19**, 1338–1348 (2010). https://doi.org/10.1109/TIP.2009.2039664

Olden, J.D., Jackson, D.A.: Illuminating the "black box'': a randomization approach for understanding variable contributions in artificial neural networks. Ecol. Model. **154**, 135–150 (2002). https://doi.org/10.1016/S0304-3800(02)00064-9

Ottomanelli, M., Caggiani, L., Iannucci, G., Sassanelli, D.: An adaptive neuro-fuzzy inference system for simulation of pedestrians behaviour at unsignalized roadway crossings. Adv. Intell. Soft Comput. **75**, 255–262 (2010). https://doi.org/10.1007/978-3-642-11282-9_27

Papatheocharous, E., Andreou, A.S.: On the problem of attribute selection for software cost estimation: input backward elimination using artificial neural networks. In: Artificial Intelligence Applications and Innovations, pp. 287–294 (2010)

Pfeiffer, M., Paolo, G., Sommer, H., Nieto, J., Siegwart, R., Cadena, C.: A data-driven model for interaction-aware pedestrian motion prediction in object cluttered environments. In: 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 5921–5928 (2018). https://doi.org/10.1109/ICRA.2018.8461157

Qiu, F., Hu, X.: Modeling group structures in pedestrian crowd simulation. Simul. Model. Pract. Theory **18**, 190–205 (2010). https://doi.org/10.1016/j.simpat.2009.10.005

Qu, Y., Xiao, Y., Wua, J., Tang, T., Gao, Z.: Modeling detour behavior of pedestrian dynamics under different conditions. Physica A **492**, 1153–1167 (2018). https://doi.org/10.1016/j.physa.2017.11.044

Robin, T., Antonini, G., Bierlaire, M., Cruz, J.: Specification, estimation and validation of a pedestrian walking behavior model. Transp. Res. B Methodol. **43**, 36–56 (2009). https://doi.org/10.1016/j.trb.2008.06.010

Ržazanskasa, P., Verikasab, A., Vibergd, P.A., Olsson, M.C.: Predicting physiological parameters in fatiguing bicycling exercises using muscle activation timing. Biomed. Signal Process. Control **35**, 19–29 (2017). https://doi.org/10.1016/j.bspc.2017.02.011

Song, X., Han, D., Sun, J., Zhang, Z.: A data-driven neural network approach to simulate pedestrian movement. Physica A **509**, 827–844 (2018). https://doi.org/10.1016/j.physa.2018.06.045

Tan, S.B., Wee, S.B., Cheung, Y.B.: Agreement or prediction: asking and answering the right question. Ann. Acad. Med. Singap. **31**(3), 405–407 (2002)

Wang, J., Li, N., Zhang, L.: Small group behaviors and their impacts on pedestrian evacuation. In: 2015 27th Chinese Control and Decision Conference (CCDC), pp. 232–237 (2015). https://doi.org/10.1109/CCDC.2015.7161696

Wang, K., Shi, X., Goh, A.P.X., Qian, S.: A machine learning based study on pedestrian movement dynamics under emergency evacuation. Fire Saf. J. **106**, 163–176 (2019). https://doi.org/10.1016/j.firesaf.2019.04.008

Wang, Z., Song, B., Qin, Y., Zhu, W., Jia, L.: Effect of vertical grouping behavior on pedestrian evacuation efficiency. Physica A **392**, 4874–4883 (2013). https://doi.org/10.1016/j.physa.2013.06.015

Xiao, Y., Gao, Z., Qu, Y., Li, X.: A pedestrian flow model considering the impact of local density: Voronoi diagram based heuristics approach. Transp. Res. C Emerg. Technol. **68**, 566–580 (2016). https://doi.org/10.1016/j.trc.2016.05.012

Xu, Y., Piao, Z., Gao, S.: Encoding crowd interaction with deep neural network for pedestrian trajectory prediction. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5275–5284 (2018). https://doi.org/10.1109/CVPR.2018.00553

Xue, H., Huynh, D.Q., Reynolds, M.: SS-LSTM: a hierarchical LSTM model for pedestrian trajectory prediction. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1186–1194 (2018). https://doi.org/10.1109/WACV.2018.00135

Yi, S., Li, H., Wang, X.: Pedestrian behavior understanding and prediction with deep neural networks. In: Computer Vision—ECCV 2016, vol. 9905, pp. 263–279 (2016). https://doi.org/10.1007/978-3-319-46448-0_16

You, L., Hu, J., Gu, M., Fan, W., Zhang, H.: The simulation and analysis of small group effect in crowd evacuation. Phys. Lett. A **380**, 3340–3348 (2016). https://doi.org/10.1016/j.physleta.2016.08.012

Yousefi, P., Naser, G., Mohammadi, H.: Surface water quality model: impacts of influential variables. J. Water Resour. Plan. Manag. **144**(5), 04018015 (2018). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000900

Yuen, J.K., Lee, E.W.M., Lam, W.: An intelligence-based route choice model for pedestrian flow in a transportation station. Appl. Soft Comput. **24**, 31–39 (2014). https://doi.org/10.1016/j.asoc.2014.05.031

Zacharias, J.: Path choice and visual stimuli: Signs of human activity and architecture. J. Environ. Psychol. **21**, 341–352 (2001). https://doi.org/10.1006/jevp.2001.0225

Zainuddin, Z., Lim, E.A.: Intelligent exit-selection behaviors during a room evacuation. Chin. Phys. Lett. **29**, 018901 (2012). https://doi.org/10.1088/0256-307X/29/1/018901

Zanlungo, F., Ikeda, T., Kanda, T.: Potential for the dynamics of pedestrians in a socially interacting group. Phys. Rev. E **89**, 012811 (2014). https://doi.org/10.1103/PhysRevE.89.012811

Zhang, J., Liu, H., Li, Y., Qin, X., Wang, S.: Video-driven group behavior simulation based on social comparison theory. Physica A **512**, 620–634 (2018). https://doi.org/10.1016/j.physa.2018.08.046

Zheng, M., Kashimori, Y., Kambara, T.: A model describing collective behaviors of pedestrians with various personalities in danger situations. In: Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02., vol. 4, pp. 2083–2087 (2002). https://doi.org/10.1109/ICONIP.2002.1199043

Shi Sun is pursuing a doctoral degree in the Architecture Department at the Harbin Institute of Technology. His research interest is in the fields of urban pedestrian space design and pedestrian behaviour modelling.

Cheng Sun, PhD, is a professor in the Architecture Department at the Harbin Institute of Technology. He is working on public building design and theory, architecture and human settlements in cold regions, urban disaster prevention and building safety, parametric architectural design, smart building and construction.

Dorine C. Duives, PhD, is an assistant professor in the Department of Transport & Planning at the Delft University of Technology. Her expertise lies in traffic operations and management, especially in the context of the active modes of transportation, such as pedestrians and cyclists, within urban environments.

Serge P. Hoogendoorn, PhD, is a professor in the Department of Transport & Planning at the Delft University of Technology. His research evolves around Smart Urban Mobility include theory, modelling, and simulation of traffic and transportation networks; development of methods for integrated management of transportation networks; impact of uncertainty of travel behaviour and network operations; impact of ICT on network flow operations, robustness and resilience; urban data and its applications.

Authors and Affiliations

Shi Sun[1] · Cheng Sun[1] · Dorine C. Duives[2] · Serge P. Hoogendoorn[2]

Cheng Sun
suncheng@hit.edu.cn

Dorine C. Duives
D.C.Duives@tudelft.nl

Serge P. Hoogendoorn
s.p.hoogendoorn@tudelft.nl

[1] School of Architecture, Harbin Institute of Technology; Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology, Harbin 150006, China

[2] Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2628 CN Delft, Netherlands