Research Article

Statistical Calculation of Dense Crowd Flow Antiobscuring Method considering Video Continuity

Huiqiang Tao

College of Mathematics and Statistics, Huanghuai University, Zhumadian 463000, China

Correspondence should be addressed to Huiqiang Tao; taohuiqiang@huanghuai.edu.cn

Received 2 February 2022; Revised 4 March 2022; Accepted 9 March 2022; Published 31 March 2022

Academic Editor: Naeem Jan

Copyright © 2022 Huiqiang Tao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

People flow statistics have important research value in areas such as intelligent security. Accurately identifying the occluded target in video surveillance is a difficulty in the video surveillance system. Now the popular moving object tracking algorithm is based on detection and cannot accurately determine the relationship between overlapping. For the statistics of people flow in the video surveillance system, a dense crowd flow antiocclusion statistical algorithm considering video continuity is proposed. This study focuses on the improved faster R-CN algorithm for small target detection, moving target correlation matching, and two-way human flow intelligent statistics. According to the small-scale characteristics of the human head target, the faster R-CNNV network structure is adaptively improved. The shallow images features are used to improve the feature extraction ability of the network for small targets. The occlusion relationship function is constructed to clearly express the relationship between the occlusion targets, and it is incorporated into the framework of the tracking algorithm. A tracking algorithm based on trajectory prediction is used to follow moving targets in real time, and a two-way human flow intelligent statistical method is used to accomplish human flow. To prove the strength of the method, tests are managed in scenes with different degrees of density, and the results show that the improved target detection algorithm improves the average accuracy of 7.31% and 10.71% on the Brainwash test set and Pets2009 benchmark data set, respectively, compared with the original algorithm. The F-value of the comprehensive evaluation index of video stream of people intelligent statistical method in various scenes can reach more than 90%. Compared with the excellent methods SSD sorting algorithm and yolov3 deepsort algorithm in recent years, its F value is increased by 1.14%–3.04%.

1. Introduction

Intelligent video surveillance equipment is commonly available in public spaces such as shopping malls, hospitals, schools, and scenic spots, among other places, and serves as a vital method of security. One of the most important research contents of intelligent video surveillance is the study of people flow statistics, which has a significant impact on the research in areas such as intelligent security, intelligent tourism, postdisaster rescue, traffic planning, and others [1–3].

In order to realize communication and interaction between the natural world and people, smart buildings make use of modern information technologies such as the internet of things, cloud computing, intelligent perception, and intelligent perception. People should always be at the center of any smart building’s service provision. According to a relevant study, “people spend 80 percent to 90 percent of their time in the indoor building environment every day, thus the perception of indoor personnel and personnel movement has become a smart building in order to deliver a better experience for indoor personnel.” This is the fundamental premise. The statistical data of people flow in the indoor office area can be used to provide reliable data support for the automatic control of temperature and humidity, light intensity, elevator planning, station planning, and other aspects of the building’s interior environment, resulting in a comfortable and humanized office environment. In the case of shopping malls, for example, we use passenger flow information to analyze customers’ purchase patterns and preferences, coordinate the layout of shopping malls, arrange the placement of goods, and reduce the waste of
commodity resources. These are examples of how to use passenger flow information [4–6]. However, in confined areas such as hallways, the lighting conditions are weak, and there are possible safety issues such as crowding and trampling that must be considered. The quantity of individuals and the direction in which they are moving may be determined in real time, allowing for early warning and the proper deployment of evacuation and drainage measures to be implemented [7, 8].

In summary, people flow statistics technology not only provides a reliable foundation for the automation and intelligent control of smart buildings but also plays an important role in the rational allocation of resources, as well as the prevention and control of public safety incidents, and it has a wide range of application prospects as well as social significance.

Prior to the development of machine learning technology, the task of counting people flow in the surveillance video was mostly done manually, which not only required a large amount of manpower and material resources but also resulted in a large amount of information in the video that could not be properly understood by the computer. In tandem with the continual advancement of computer vision and machine learning technology, methods of people flow statistics based on shallow learning are continuing to emerge as viable alternatives. Given the artificial nature of the design and feature extraction employed by such methods, feature redundancy is substantial, and meaningful information cannot be recovered in an intelligent manner, resulting in low statistical accuracy. A deep neural network may be built to efficiently extract deep picture features through automatic learning, which can then be used to solve the difficulties mentioned above [9–15].

Methods for calculating people flow statistics based on deep learning have been proposed on a consistent basis since the dawn of the era of artificial intelligence. The most significant distinction between it and traditional methods is that it can automatically learn features from data, without the need for manual feature creation and extraction, as well as without the need for foreground segmentation. Cross-scene counting has been accomplished by some researchers using convolutional neural network models to generate density maps, which are more efficient and robust than hand-designed features in terms of cross-scene counting. However, even though this type of method does not require the detection of each pedestrian, the number of people can be obtained directly by regression or integration of the crowd density map; however, it is precise because the specific information about each pedestrian cannot be obtained that subsequent pedestrian trajectory analysis is rendered impractical research. People flow statistics methods based on deep learning detection have also been proposed in recent years, in part as a result of the ongoing breakthroughs made by deep learning in the field of target identification. The use of a convolutional neural network to analyze video pedestrian traffic has been proposed by several researchers. While this method captures the statistics of people’s movement in the monitoring region, it may result in missed detections when pedestrians obstruct each other. With the help of the SSD (single shot multibox detector) and KCF (kernel correlation filter) algorithms, Zhang discovers and tracks human head targets and then analyzes the target trajectory to attain bidirectional counting [16, 17]. This method effectively solves the problem of occlusion between pedestrians by detecting human head targets. However, because this method employs the strategy of identifying the pedestrian’s head from precisely above the pedestrian, the algorithm’s applicability situations are severely restricted. Due to differences in poses, scales, and occlusions of pedestrians in the video, the present approaches have poor accuracy in estimating people flow statistics across diverse circumstances, despite the fact that the methods described above have achieved certain outcomes.

In order to address the issues raised above, the purpose of this work is to conduct research on intelligent statistical approaches for dense crowd movement antiocclusion while taking video continuity into consideration. We expect to achieve accurate and intelligent statistics of video traffic by investigating the improved faster R-CNN algorithm for small target detection, discussing target tracking technology that is based on trajectory prediction, designing a two-way human traffic intelligent statistics algorithm, and constructing an experimental environment for experimental analysis [18–20].

The paper arrangements are as follows:

Section 2 examines the proposed method and discusses different subsections. They demonstrate the multitarget tracking module and two-way people counting module based on three-dimensional center of mass. Section 3 explains and compares the different data experiments. Section 4 concludes the article.

2. The Proposed Method

This paper proposes the overall architecture of the two-way people flow counting method fused with depth information. The algorithm architecture steps are as follows:

1. RGBD video sequence extraction module. The Kinect camera is used to shoot RGB images and depth images of a particular place at the same time, and the RGB images are paired with the depth images in the same incident to form a set of RGBD images and output RGBD video sequences.

2. YOLOv3 person detection module. Taking the RGB video sequence obtained in the above step (1) as input, the YOLOv3 target detection algorithm is used to detect the pedestrian head, and the frame coordinates of the head area of all persons in the current frame of the RGB video sequence are obtained and output.

3. A multitarget tracking module fused with depth information. The coordinates of the pedestrian head frame obtained in step (2) are used as input and marked as the starting position of the pedestrian trajectory, and the pedestrian is tracked. The depth
change rate and frame IOU are used to optimize the target matching process, and the context depth difference judgment is combined. Occlusion status adaptively adjusts the update, logs out of the trajectory, optimizes the trajectory processing process, and outputs the successfully matched pedestrian trajectory and occlusion status information.

(4) Four-dimensional center-of-mass coordinates are used in the bidirectional pedestrian flow statistics module. It receives as input the pedestrian trajectory and blocking state obtained, determines the movement direction of people in the current frame based on the trajectory blocking state and 3D center of mass coordinates, counts the people passing through the main counting line in both directions according to their movement direction and trajectory, and outputs the counting result, people blocking state, people movement trajectory, and horizontal and vertical movement means.

(5) Visualization of information on people’s movements. It is possible to count and visualize the number of individuals present, as well as to view a real-time log, the trajectory of people’s movement, the state of people’s blocking, and the horizontal and vertical movement directions. RGB and depth image grouping in video sequences is shown in Figure 1. The test result is shown in Figure 2.

Figures 1 and 2 are not merged together because Figure 1 depicts the grouping of RGB and depth images in video sequences and Figure 2 shows the result of every person.

2.1. Multitarget Tracking Module with Fused Depth Information. This paper constructs on the Deep SORT algorithm by incorporating incident depth information and introducing two new steps: “target matching strategy based on the joint constraint of IOU and depth change rate” and “determining the occlusion incident by using the depth difference of the trajectory context,” which both improve the accuracy and reliability of tracking matching while maintaining the same level of achievement. Figure 3 depicts the entire flow of the procedure. Using the depth change rate as a constraint on the target trajectory, it is possible to improve the reliability of target matching. At the same time, the module determines whether or not a target is occluded by combining the depth difference of the target trajectory context in order to optimize the target trajectory.

The target matching strategy based on the joint constraint of border IOU and depth rate of change provides input data for the target matching strategy based on the joint constraint of border IOU and depth rate of change after extracting the appearance description of the target using the pre-trained CNN network in the Deep SORT algorithm and obtaining two motion information of the target by the Kalman filter prediction: the position information and the uncertainty information. The structure of pretrained CNN is shown in Figure 4.

In the classic method, hard spacing is used to categorize IOU, and hard spacing entails using a hyperplane to correctly classify all values. Nevertheless, the reliability and consistency of IOU’s hard interval are low, and it is prone to match errors. In this study, we present a joint constrained target matching approach that uses a soft interval to smooth the decision boundary, that is, that incorporates the border IOU and depth rate of change as slack quantities, which allows for a certain amount of classification error. This rate is defined as follows: the average depth value of the 1st frame of the trajectory storage minus the average depth of the target in the current i-th frame, and then divided by 100 to get the percentage, and the calculation formula is shown in equation (1). The depth change rate \( dr \) is defined as follows:

\[
\begin{align*}
dr &= \min \left( \frac{D_{trk} - D_{det}}{100} \right).
\end{align*}
\]  

The specific steps are as follows:

(1) Calculate the IOU cost matrix. Calculate the IOU cost matrix \( C \) between the set \( trk \) of all unmatched trajectories in the initial matching result and the set \( det \) of unmatched detections. IOU cost can be obtained by the following equations:

\[
\begin{align*}
\text{iou}_{trk}^{det} &= \frac{\text{bbox}_{det} \cap \text{bbox}_{trk}}{\text{bbox}_{det} \cup \text{bbox}_{trk}},
\end{align*}
\] 

\[
\begin{align*}
C_{ij} &= 1 - \text{iou}_{det}^{trk},
\end{align*}
\] 

where \( trk \in trk (i = 1, 2, \ldots, M) \) is the i-th episode of set, \( det \in det (j = 1, 2, \ldots, N) \) is the j-th episode of set, and \( \text{bbox}_{det} \) and \( \text{bbox}_{trk} \) are the detection edges and the trajectory-preserving edges, respectively. A larger value of \( \text{iou}_{det}^{trk} \) indicates a higher position overlap and a smaller IOU valuation, and vice versa. IOU valuation describes the position similarity of the target motion between frames and takes the value range \([0, 1]\).

(2) Calculate the depth change rate cost matrix \( dr \). Obtain the depth mean value \( D_{trk} \) of all unmatched trajectories and the depth mean value \( D_{det} \) of all unmatched detection results and calculate the depth change rate matrix \( dr \) between two, which is calculated as follows:

\[
\begin{align*}
dr_{ij} &= \min \left( \frac{|D_{trk} - D_{det}|}{100}, 1 \right).
\end{align*}
\] 

(3) Joint constraint the decision boundary is smoothed by using a soft interval: if \( 0.6 \leq C_{ij} \leq 0.8 \), hen we use \( C_{ij} \) and \( dr_{ij} \) to joint constraint cost matrix if \( C_{ij} > 0.8 \), it indicates that the location similarity of the two is very low, \( dr_{ij} \geq 0.8 \), it indicates that the depth similarity of the two is very small, and the above two cases reject \( trk \) and \( det \) if \( 0 \leq C_{ij} \leq 0.6 \), it indicates that the location similarity of the two has high reliability. The joint constraint formula is as follows:
Figure 1: RGB and depth image grouping in video sequences.

Figure 2: Test result.

Figure 3: Framework of statistical calculation of dense crowd flow antiobscuring method considering video continuity.
The specific algorithm is as follows:

(1) Define the trajectory attributes. For each initialized trajectory, three parameters are assigned: occState, which keeps track of the current trajectory's occlusion state, and unoccTimer, which keeps track of how long the current trajectory is unobstructed; the average depth value of the current detection result or trajectory is roiDepthMean.

(2) To calculate the depth difference $D$ between the trajectory and the detection result, firstly, we need to calculate the average depth of the two ROIs and then map the target boundary coordinates to the depth image to extract the human head region; in this paper, we fill the depth missing value by calculating the minimum value of all nonzero elements in the region of interest to reduce the interference of background pixel depth. Furthermore, a rectangular window of size $6 \times 6$ is expanded outward with the center point of the region of interest as the center, and the depth mean under this rectangular window is calculated as the depth mean of the region of interest. Finally, the depth mean value $D_{\text{det}}$ of the detection result at the current moment and the depth mean value $D_{\text{trk}}$ stored in the corresponding trajectory at the previous moment are obtained, and the depth difference between frames is calculated by equation (9) for subsequent occlusion judgment.

(3) Determine the occlusion status of the target using the trajectory contextual depth difference. Based on the correlation between occlusion change and depth change, this paper uses the trajectory contextual depth difference to determine the occlusion status of the target, and when the depth difference between the current frame and the previous frame is too large,
the target is judged to be occluded, and when the target is not occluded for \( N \) consecutive times, the target is judged to be unoccluded.

The correlation between occlusion change and depth change is analyzed. Due to the continuity of human movement in the tracking process, the depth difference \( D \) between frames will be kept within a small value when the head is in the unobstructed state. When the head is obscured by a solid object, the corresponding \( D \) changes as follows: value fluctuation – small value – value fluctuation. When the head is blocked by a hollow object, the corresponding \( D \) will fluctuate between small and large values within a certain period of time. When the head is overlapped by multiple people, in this case, because the movement direction and speed of the people are usually random, the situation is more complicated, and it is difficult to capture the significant change of \( D \) quantitatively.

The occlusion status is judged by the depth difference of the trajectory context. When the interframe depth difference \( D \) is greater than or equal to the maximum depth difference threshold \( T \), the occlusion status of the current target is updated to “occluded”; if the interframe depth difference of the trajectory in \( N \) consecutive frames meets \( D < T \), the occlusion status of the current target is updated to “unoccluded.” In this paper, the value of \( N \) is 5, and the binary formula of target occlusion status is as follows:

\[
F(\text{occState}) = \begin{cases} 
\text{uncovered, occState} = 0, \\
\text{covered, occState} = 1. 
\end{cases} 
\]  

(7)

(4) The trajectory processing strategy is optimized based on the occlusion state. For the unmatched trajectory, this paper adaptively adjusts the maximum survival time max-ages of the trajectory by using the occlusion state of the target. First, we set a fixed initial value \( S \) (15 in this paper) for max-ages; if the number of disappearing frames of the current trajectory \( \text{time since update} \leq S \), the trajectory is retained for subsequent matching; otherwise, the occlusion state of the trajectory is checked. If the current trajectory is in “no occlusion” state, there is enough evidence that the target corresponding to the trajectory has disappeared from the video frame, so the trajectory is deleted; if the trajectory is in “occluded state”, make max-ages + 1 and extend the maximum survival time of the trajectory appropriately. At the same time, set the maximum cutoff value end and delete the track when the track in the “blocked” state meets \( \text{max-ages} \geq \text{end} \), so as to avoid storing too many invalid tracks in the tracking process.

2.2. Two-Way People Counting Module Based on Three-Dimensional Center-of-Mass Coordinates. First, in view of the certainty that 2D center of mass coordinates is easy to judge the direction of people’s movement in this paper, to judge the direction of movement, depth information is introduced, and the target’s 3D center of mass coordinates are used, thereby improving the accuracy of judging the direction of target movement; second, the two-way people counting method based on the 3D center of mass coordinates is introduced and used to improve the accuracy of judging the direction of target movement.

Because the traditional method only considers changes in the two-dimensional coordinates of the target trajectory between two frames in order to determine the direction of movement of the personnel, it is easy to make a mistake when judging the direction of movement of irregularly moving personnel crossing the line. Furthermore, because the traditional technique only evaluates changes in the two-dimensional plane coordinate system and ignores changes in the three-dimensional distance changes of the target in the picture, it is difficult to make a mistake when determining the direction of movement of the object. We use the 3D center of mass coordinates to determine the lateral and vertical motion direction of a person. The depth information of the target is introduced, and multiple frames of the center of mass are combined at the same time in order to determine the lateral and vertical motion direction of the person; the specific steps of the algorithm are as follows:

(1) Get the 3D center of mass coordinates \( \text{Centroid}_{\text{det}} \). For the trajectory that matches successfully, get the 3D center of mass coordinates \( \text{Centroid}_{\text{det}} \) of the current frame detection result, the 3D center of mass can be expressed as \( (x, y, d) \), where \( x, y, d \) are the center of mass horizontal coordinates, vertical coordinates, and depth of mass, respectively.

Each trajectory has a \( \text{preCentroid} \) queue of maximum length \( N \), \( \text{preCentroids} \), which is used to record the 3D center of mass coordinates of the person in the previous \( N \) frames. In this paper, \( N \) is taken as 10, that is, the maximum trajectory can record the 3D centroids successfully matched 10 times from the current moment forward.

(3) Determine the direction of movement of the person. The method proposed in this paper can realize the direction judgment of vertical (up and down) and horizontal (left and right) simultaneously.

For the longitudinal direction judgment, there are two cases as follows. If the person is in the unobstructed state at the current moment, the difference between the mean value of the depth of the center of mass of the previous IV frames and the current depth of the center of mass is calculated to obtain the direction of the target movement, as shown in the following equation:

\[
diff_d = \text{mean}\left(\text{preCentroids}^d \right) - \text{centroid}_{\text{det}}^d. 
\]  

(8)

where \( \text{preCentroids}^d = [d_1, d_2, \ldots, d_N] \).

If the person is obscured at the current moment, the depth of the center of mass fluctuates and is not suitable for quantitative calculation, so the mean value of the longitudinal coordinates of the center of mass of the first IV frames
iscalculatedandthecurrentlongitudinalcoordinatesofthe
centerofmassaremadethedifferencetogetthedirectionof
the target movement, as shown in the following equation:
\[ \text{diff}_y = \text{centroid}_y^t - \text{mean}(\text{preCentroids}_y), \]
(9)
where \( \text{preCentroids}_y = [y_1, y_2, \ldots, y_N] \).

Then, based on the sign of the target occlusion state and
\( \text{diff} \), the longitudinal movement direction of the person
at the current moment is judged by using the following
equations:
\[ \text{diff} = \begin{cases} 
  \text{diff}_d, & \text{if occState} = 0 \\
  \text{diff}_y, & \text{if occState} = 1
\end{cases} \]
(10)
\[ \text{diff} = \begin{cases} 
  \text{up}, & \text{if diff} = 0 \\
  \text{down}, & \text{if diff} = 1 \\
  0, & \text{if diff} = 1.
\end{cases} \]
(11)

Finally, update the trajectory parameters. After finishing
the direction judgment of the personnel in the current frame,
we use the calculation result of the current frame to update
the parameters of 3D center of mass coordinates, horizontal/
vertical motion direction, and average depth value of the
trajectory and press the 3D center of mass coordinates of
the current frame into the pre-Centroids queue for the judgment
and processing of the subsequent frames.

3. Experimental Analysis and Comparison

3.1. Data Set Selection. In the actual scene, the longitudinal
counting along the camera lens normal movement is the most
widely used, and the horizontal counting is basically the same
as its principle. The MICC data set is provided by Bondi et al.
The MICC data set is a dense scene with a 45° pitch angle,
which consists of three sets of FLOW, QUEUE, and GROUPS.
The scene schematic is shown in Figure 5, which contains a
variety of pedestrian movement patterns, which is challenging,
and the three data sets are described in detail below.

As illustrated in Figure 5(a), the FLOW data set has 1,148
frames, each of which contains 9 individuals. People moving
quickly in underground corridors and via building entrances
and exits are simulated, as are some lateral movements in the
simulation. This sequence is intended to evaluate the sta-
tistical accuracy of the suggested method when it is subjected
to rapid movements and sudden changes in directions.

The GROUP data set, depicted in Figure 5(b), contains 917
frames, each of which contains one person. The participants
enter the frame and begin to converse with one another in two
groups, simulating the progression of the action from move-
ment to gathering and remaining in the conference room,
among other things. These steps are used to test the accuracy
of the suggested algorithm for people counting in the context of a
crowd gathering, which is the subject of this study.

As depicted in Figure 5(c), the QUEUE data set consists
of 1,128 frames, each of which contains 10 individuals. The
people move steadily ahead in an organized queue, simulat-
ing the queuing scenarios found in stores and ticket halls,
as the music plays in the background. For the purposes of
this research, we will utilize this sequence to check that the
suggested algorithm for people counting is accurate in the
situation of modest movements of individuals.

3.2. Evaluation Metrics. The performance of our algorithm is
reflected by the precision, recall, and \( F \).
\[
\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}
\]
\[
\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}
\]
\[
F = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}
\]
(12)
In order to quantitatively evaluate the performance of the intelligent statistical method of video foot traffic in this paper, the number of positive checks, false checks, and missed checks in the test videos were manually counted, and the recall rate, accuracy rate, and $F$-value were calculated from the above statistics, and the results are shown in Table 1. From the results, it can be seen that although the number of false detections in the intensive scenes leads to a lower precision rate criterion than the normal scenes, the overall performance of the method in the three videos is better in terms of recall rate, precision rate, and $F$-value criterion, which all reach more than 90%.

The results of the pedestrian flow statistics test are shown in Figure 6, in which the red box is the detection result, the white box is the tracking result, the yellow text in the upper left corner is the real-time statistics result of pedestrian flow, and the area surrounded by two yellow virtual counting lines is the target area. From the figure, it can be seen that this method can detect and track all the pedestrians and make accurate statistics of pedestrian flow. In addition, it can also achieve high accuracy detection and tracking in the case of dense and partially overlapping pedestrians. It can be proved that this method has good accuracy and certain robustness.

4. Conclusion

(1) In order to solve the problem of poor reliability of association strategy based on IOU distance in dense crushes, this paper proposes the statistical calculation method of dense crowd traffic anti-occlusion considering video continuity, which uses location similarity and depth similarity to jointly constrain the target and motion trajectory to improve the matching reliability.

(2) Propose a target occlusion judgment technique based on the depth difference of the trajectory context in order to accurately assess the target occlusion status and adopt the trajectory update strategy in real time.

(3) Since the irregular motion of people and the motion along the imaging plane normal has a small difference in two-dimensional coordinates, this paper proposes to use three-dimensional center-of-mass coordinates to determine the direction of motion and introduce depth information to get the location difference between targets so that the above problems can be solved.

At this stage, the technology only allows for the location and tracking of people in a two-dimensional plane; however, the next step will be to locate people in three-dimensional space and study their distribution. Moreover, the current method can carry a maximum of about 10–15 people, and the subsequent attention will optimize the method to improve the capacity of carrying people.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.
Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] J Desyllas, E Duxbury, J Ward, and A. Smith, “Pedestrian Demand Modelling of Large Cities: An Applied Example from London,” Journal of Centre for Advanced Spatial Analysis, vol. 62, no. 3, pp. 1–18, 2003.

[2] G Stavroulaki, D Bolin, M Berghauser Pont, L. Marcus, and E. Häkansson, “Statistical modelling and analysis of big data on pedestrian movement,” in Proceedings of the 12th International Space Syntax Symposium, vol. 79, pp. 71–79, Beijing, China, 2019.

[3] J. Ye, X. Chen, C. Yang, and J. Wu, “Walking behavior and pedestrian flow characteristics for different types of walking facilities,” Transportation Research Record: Journal of the Transportation Research Board, vol. 2048, no. 1, pp. 43–51, 2008.

[4] W. H. K. Lam, J. Y. S. Lee, and C. Y. Cheung. “A study of the bi-directional pedestrian flow characteristics at Hong Kong signalized crosswalk facilities,” Transportation, vol. 29, no. 2, pp. 169–192, 2002.

[5] M. R Virkler and S Elayadath, “Pedestrian speed-flow-density relationships,” Journal of Videos, vol. 12, no. 1, pp. 112–119, 1994.

[6] S. C. Wong, W. L. Leung, S. H. Chan et al., “Bidirectional pedestrian stream model with oblique intersecting angle,” Journal of Transportation Engineering, vol. 136, no. 3, pp. 234–242, 2010.

[7] J. A. Vallejo-Borda, H. A. Ortiz-Ramirez, A. Rodriguez-Valencia, and J. Ortizur, “Forecasting the quality of service of bogota’s sidewalks from pedestrian perceptions: an ordered probit MIMIC approach,” Transportation Research Record: Journal of the Transportation Research Board, vol. 2674, no. 1, pp. 205–216, 2020.

[8] N Raford and D Ragland, “Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis,” The Case of Boston Massachusetts, vol. 32, 2006.

[9] J Huo, X Fu, Z Liu, and Q. Zhang, “Short-term Estimation and Prediction of Pedestrian Density in Urban Hot Spots Based on mobile Phone data,” IEEE Transactions on Intelligent Transportation Systems, vol. 34, 2021.

[10] P. Das, M. Parida, and V. K. Katiyar, “Analysis of interrelationship between pedestrian flow parameters using artificial neural network,” Journal of Modern Transportation, vol. 23, no. 4, pp. 298–309, 2015.

[11] M. Patra, E. Sala, and K. V. R. Ravishankar, “Evaluation of pedestrian flow characteristics across different facilities inside a railway station,” Transportation research procedia, vol. 25, pp. 4763–4770, 2017.

[12] S Wang, R Zhou, and L Zhao, “Forecasting Beijing transportation hub area’s pedestrian flow using modular neural network,” Discrete Dynamics in Nature and Society, vol. 2015, 2015.

[13] K. Ghasemlou, M. M. Aydin, and M. Sinan Yıldırım, “Prediction of pedal cyclists and pedestrian fatalities from total monthly accidents and registered private car numbers,” Archives of Transport, vol. 34, no. 2, pp. 29–35, 2015.

[14] L Huang, T Chen, Y Wang, and H. Yuan, “Forecasting daily pedestrian flows in the tiananmen square based on historical data and weather conditions,” ISCRAM, vol. 18, pp. 1315–1320, 2015.

[15] X. Chen, J. Ye, and N. Jian, “Relationships and characteristics of pedestrian traffic flow in confined passageways,” Transportation Research Record: Journal of the Transportation Research Board, vol. 2198, no. 1, pp. 32–40, 2010.

[16] A. Zhen, Q. Wang, A. Wang, and K. Wu, “Design and implementation of pedestrian flow monitoring and warning platform in public area based on pedestrian detection,” Journal of Physics: Conference Series, vol. 1617, no. 1, Article ID 012074, 2020.

[17] D. A. Obeng and M Salifu, “Modelling risk factors of pedestrian accidents on trunk roads in Ghana,” International Refereed Journal of Engineering and Science (IRJES), vol. 2, no. 5, pp. 55–64, 2013.

[18] C. Feliciani, A. Gorrini, L. Crociani, G. Vizzari, K. Nishinari, and S. Bandini, “Calibration and validation of a simulation model for predicting pedestrian fatalities at unsignalized crosswalks by means of statistical traffic data,” Journal of Traffic and Transportation Engineering, vol. 7, no. 1, pp. 1–18, 2020.

[19] C. Mao, L. Bao, S. Yang, W. Xu, and Q. Wang, “Analysis and prediction of pedestrians’ violation behavior at the intersection based on a Markov chain,” Sustainability, vol. 13, no. 10, p. 5690, 2021.

[20] L. M. Pfiester, R. G. Thompson, and L. Zhang, “Spatiotemporal exploration of Melbourne pedestrian demand,” Journal of Transport Geography, vol. 95, Article ID 103151, 2021.