ISSN 1870-4069

Analysis of Walking Paths from Pedestrian Tracking in Real-Time Using Deep Learning

Ana L. Ballinas-Hernández[1], Carlos E. Hernández-Inzunza[2], M. Claudia Denicia-Carral[1], Maricruz Rangel-Galván[3]

[1] Benemérita Universidad Autónoma de Puebla,
Complejo Regional Centro,
Mexico

[2] Universidad Autónoma de Sinaloa,
Facultad de Informatica,
Mexico

[3] Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias Químicas,
Mexico

analuisa.ballinas@correo.buap.mx, carloshinzunza2@gmail.com,
claudia.denicia@correo.buap.mx, maricruz.rangelgalvan@viep.com.mx

Abstract. The study of pedestrian walking has a crucial role in design of safe and comfortable public spaces in urban areas and in smart cities. Analysis of walking paths is a task with great challenges in various fields to understand and characterize pedestrian behavior. In this work, real-time detection and tracking algorithms of pedestrians are applied to recover walking paths in urban environments. The main contribution of this work is the recovery of microscopic parameters and frequency distribution analysis of average speed to evaluate the counterflow pedestrian walking dynamics, as well as the description of paths generated from real experiments.

Keywords: Pedestrian detection, pedestrian tracking, deep learning, walking paths.

1 Introduction

Pedestrian walking dynamics has been a subject of wide interest in recent years and is currently an open field of study. The importance of analyzing this system is mainly oriented towards safe and comfortable design of urban spaces, as well as the construction of pedestrian zones with adequate infrastructure in smart cities.

Pedestrian detection and tracking has various applications in the field of computer vision, such as autonomous driving, driver assistance, video surveillance systems, robotics, among others [8]. Traditional object tracking techniques can be inefficient and unreliable, especially in challenging scenarios where there are changes in lighting in the environment, partial obstructions, variation of postures, areas with high density of crowds, etc. Several algorithms have been developed to solve these situations, however, many problems remain to be solved.

Walking styles vary in each geographic region because pedestrians take into account the cultural aspects of their environment, their physical characteristics, and their walking preferences. It is important to consider these aspects to build reliable pedestrian detection and tracking systems in different scenarios.

In the current work, the detection and monitoring of pedestrians is done in real time, in regular walking conditions, considering different environments, people with different physical characteristics, different walking preferences and at low, medium and high densities. Walking paths and the microscopic parameters of each pedestrian with respect to their position and speed are recovered. In addition, walking patterns and the average speed obtained by following pedestrians are analyzed to understand the group behavior.

2 Related Work

In the work of Camara et al. a review of the current state of the art in the field of detection, recognition, monitoring and prediction of pedestrian paths has been done [3]. Review is organized into five levels, the lower levels being pedestrian detection methods that rely on machine vision and robotics models to detect pedestrians, tracking their positions and speeds over time. At the highest levels, other factors intervene, among them psychological and pedestrian personality, through which it is possible to predict their movements and actions. At these levels, psychological information is inferred from body language, gestures, and demographic information. Models like YOLO (You Only Look Once) have been widely used in object detection techniques.

In the work of Sundararaman et al. a head tracker in high densities comparable to traditional pedestrian trackers is introduced. Results show their method is comparable with traditional algorithms [14]. Also they present a metric known as EDEucl, two methods for head detection, and a useful model for crowd counting and movement analysis. A work has been done for the tracking of multiple objects by means of the detection and tracking of objects in a scene through a SORT algorithm (simple online real-time tracking) that includes an identification module [1]. Experiments were done with the MOT17 and MOT20 sets, relevant data sets in the field of study. Results obtained show that SORT-based algorithms for pedestrian tracking are a good option and can be easily integrated into other tracking trackers for detection.

Video surveillance systems have been developed to detect unusual events through pedestrian monitoring in order to maintain crowd safety [6]. A comparison is made between different computational vision techniques for the detection and tracking of pedestrians and some pedestrian video databases obtained from different repositories are described. In addition, the main problems that occur in human detection and tracking are identified, such as occlusion, variation in postures and areas of high crowd density that generate errors.

An occlusion management strategy has been developed through modeling the relationships between occlusions and occluded tracks, unlike traditional feature-based approaches [13]. This strategy is used for multiple pedestrian detection, focuses on lane management and is capable of working for bidirectional tracking with results higher than those reported for the MOT17, MOT20 and MOT16 sets.

Several experimental works have been done to understand real pedestrian walking dynamics through analysis of their paths and quantification of the system under controlled environmental conditions in different walking situations [11, 17]. In the work of Zanlungo et al. the effect of real pedestrian walking in groups with bidirectional flow was analyzed where a line grouping algorithm was built for its characterization [16]. The experimental results of the physical system were compared with the simulation model where the speed of pedestrians, the number of collisions, the number of sidewalks and the ratio of pedestrians are quantified.

In the work of Feliciani et al. controlled experiments of pedestrian walking in chaotic scenarios were done [5]. Authors analyzed paths, collision avoidance mechanisms and the fundamental diagram for different experimental data. From description of experiments, a simulation model based on particle gases was built where the interactions between pedestrians were modeled as physical forces. Authors concluded that when people walk in chaotic conditions with minimal influence from the environment, a simple model is sufficient to describe the system general behavior.

In present work, pedestrian walking paths recovered from videos of real experiments are analyzed. For pedestrian detection, a YOLOv3 algorithm based on convolutional neural networks and SORT algorithm for pedestrian tracking are applied [18]. The contribution of this work, with respect to the related work, is the recovery of microscopic parameters of pedestrians that characterize their walking and analysis of generated paths and average speed of the group.

3 Proposed Methodology

To generate walking paths that pedestrians follow, the methodology shown in Figure 1 is proposed. The detection and monitoring of pedestrian walking in corridors under regular situations in corridors is proposed. Some microscopic parameters of pedestrians are identified and the group behavior patterns obtained are analyzed.

3.1 Video Collection

In this phase, a set of videos of urban streets where pedestrians walk at different densities is built. Data is obtained from the public repository Caltech Pedestrian Dataset of people who walk under regular conditions with a total of 10 hours of video with a resolution of 640 x 480 px [4]. In addition, videos of pedestrians walking against the flow were captured with a total of 5,100 seconds of video in

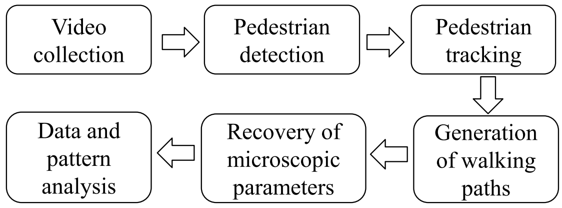


Fig. 1. Proposed methodology for trajectory analysis of pedestrian walking.

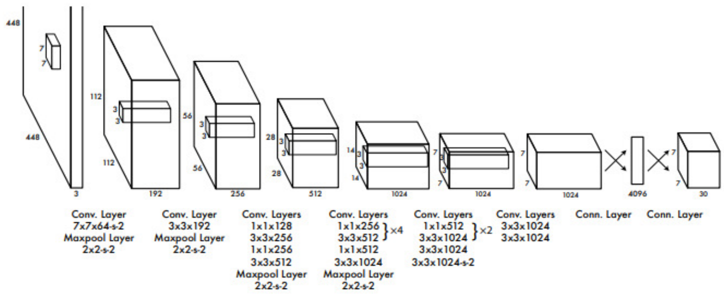


Fig. 2. YOLO architecture for object detection based on convolutional neural networks, Obtained from [10].

25 different corridors with a resolution of 1,280 x 720 px. Videos were captured during the day at different times with low, medium and high pedestrian densities.

3.2 Pedestrian Detection

Videos are divided into frames for the recognition of pedestrians in each image. The pretrained model YOLO is used for the fast detection of multiple objects that uses a CNN (convolutional neural network) following the architecture of Figure 2 [12]. The CNN has 24 convolutional layers and 2 fully connected layers from a set of images rescaled to a size of 448×448. For automatic detection and classification of pedestrians, the pretrained model YOLOv3 is used, which generates feature maps at three different scales. The algorithm predicts bounding boxes for each image cell to recognize objects and calculates the loss function that is compared with the ground truth where the highest IoU (Intersection of Union) is chosen to recognize pedestrians [18].

3.3 Pedestrian Tracking

For tracking by pedestrian detection in consecutive frames from videos, SORT algorithm is applied [2]. The algorithm combines object detection with Hungarian algorithms and Kalman filters for tracking. YOLOv3 and SORT algorithms were tested overall with 4,327 frames with 6 different classes and obtained an average accuracy of 0.724, an average precision of 0.8068, and a recall of 0.909. These algorithms allow pedestrian dynamics to be tracked in real time by assigning an ID to each pedestrian.

3.4 Generation of Walking Paths

Based on pedestrian tracking between input video frames, walking paths of each pedestrian are automatically retrieved in a new image. Each path is painted in a color for each pedestrian for easy identification. These paths allow to see the evolution of pedestrian walking from an initial position to a final position where pedestrians walk against the flow in corridors reaching the exit.

3.5 Recovery of Microscopic Parameters

Some parameters are recovered at microscopic level where pedestrians are treated as individuals. These parameters are positions (x, y) of each pedestrian, taking reference system the size of videos. In addition, average speed of each pedestrian v from the first to the last frame is retrieved. Speed relates distance travelled by pedestrians to the time elapsed as a scalar quantity. These data are stored in .CSV files. Speed is a relevant parameter to understand the way pedestrian crowds walk at different densities and under different situations. The parameters obtained are:

$$P = \{x, y, v\}. \tag{1}$$

3.6 Data and Pattern Analysis

Final phase of the methodology consists in analyzing the microscopic parameters recovered by means of a frequency distribution histogram of walking speed to quantitatively analyze the pedestrian speed behavior at different densities. In addition, obtained paths are analyzed qualitatively to understand the evolution of the group behavior on a macroscopic level. The objective is to recognize patterns of self-organization that appear in walking dynamics of pedestrian crowd.

4 Results

Results obtained in this work are: construction of a set of videos of corridors, where pedestrian crowds walk against flow under regular flow conditions; path generation obtained by detection and monitoring of pedestrians in real time; qualitative description of walking paths as a manifestation of the group self-organization on a macroscopic level; data retrieval from microscopic parameters of crowd dynamics. Three different crowd density areas are tested: low, medium, and high.

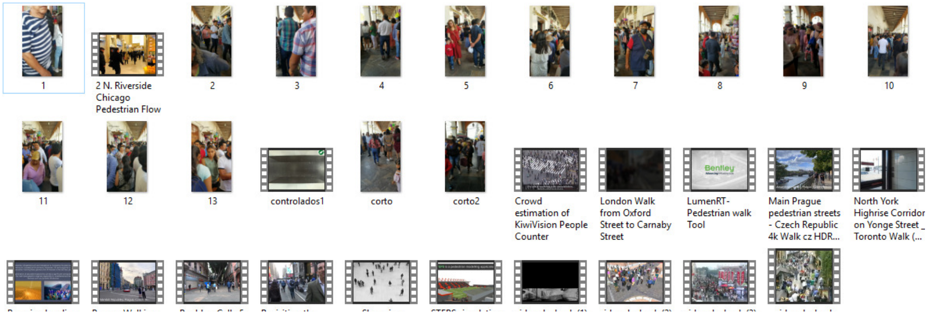


Fig. 3. Sample of the dataset that contains public repositories and own videos of pedestrians walking.

Table 1. Evaluation of detection based pedestrian tracking model.

Video	**Total frames**	**Pre-processing time**	**Inference time**	**Average ground truth**
Low density	197	0.7ms	42ms	0.7433
Medium density	310	1.1ms	42.5ms	0.7633
High density	156	1.2ms	53.8ms	0.5959

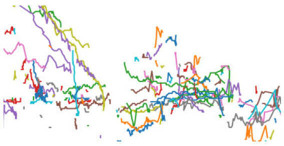
Total of 41,100 seconds of videos of pedestrians walking against the flow in corridors under regular conditions were collected (see a sample in Figure 3). Videos are divided into frames to process the image sequences to be processed.

Tests of tracking algorithms for pedestrian detection are done using a computer with an AMD Core i5 processor at 2.2GHz with 12 GB of RAM and Windows operating system. Python programming language and machine learning libraries are used to test the predictive models. For the tests, a sample of three videos each with 350 frames is considered. Left section of Figure 4 shows results of the YOLO model application for the pedestrian detection where a bounding box is drawn with the ID of each pedestrian at three densities: low, medium, and high. Right section of Figure 4 shows results of the SORT algorithm application for pedestrian tracking, where walking paths are generated from the initial video frame to the final frame. As can be seen in the figure, it was possible to recover and draw the pedestrian walking paths in videos of real experiments based on tracking by pedestrian detection at low, medium and high densities.

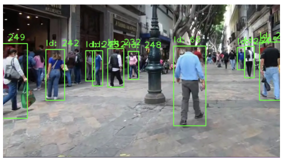
Table 1 shows average pre-processing times required for video frames at low, medium and high densities. In addition, the average ground truth of each video is calculated to evaluate results of predictive model with respect to real data.

Results of the pedestrian detection tracking demonstrate an acceptable performance of algorithms with an average ground truth of 0.7433 at low concentration densities and 0.7633 at medium densities. However, at high densities, the algorithms present a ground truth of 0.5959, which is unreliable due to the frequent occlusions of large crowds and because the separation between pedestrians is limited, causing algorithms have high detection errors.

(a) Pedestrian detection at low density



(b) Generation of walking paths at low density



(c) Pedestrian detection at medium density

(d) Generation of walking paths at medium density

(e) Pedestrian detection at high density

(f) Generation of walking paths at high density

Fig. 4. Generation of walking paths obtained by tracking pedestrians in real time.

X	204	202	201	201	200	200	199	197	194	192	191	190	189	187	185	183	181	179
Y	200	199	199	199	199	200	199	199	200	200	199	199	199	199	198	199	199	198
V	1.46	1.94	0.47	1.41	1.31	1.12	1.47	1.2	1.29	1.24	1.95	2.36	1.2	1.45	2.51	0.75	2.16	0.96

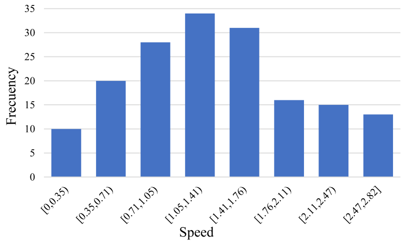
X	178	176	176	175	175	176	167	166	165	164	157	157	156	156	156	156	72
Y	198	198	198	200	200	200	199	200	199	199	198	199	199	199	199	200	207
V	1.51	1.2	2.03	2.63	0.67	1.03	2.59	1.41	2.24	0.63	1.61	0.71	1	1.52	0.41	1.01	2.24

Fig. 5. Position (x, y) and speed v obtained from pedestrian tracking.

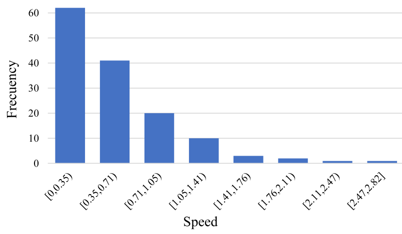
Figure 4.b shows walking paths obtained from a video at low densities. These paths present irregular patterns because, as there are few encounters, pedestrians apply turns for convenience and their walking is free. Free walking is a very commonly reported rule in the literature at low densities [9]. Figure 4.d shows paths of bidirectional pedestrian walking where line formation occurs, which is a very frequent phenomenon in the simulation of pedestrian walking as a manifestation of self-organization [7]. Figure 4.f shows the pedestrian walking paths at high densities. Despite the errors in crowd detection, it can be seen that walking lines are regular because pedestrians are not free to turn for convenience and follow a line until they reach their goal. From this description of paths, it is possible to qualitatively understand the pedestrian crowd dynamics at three different densities under regular walking conditions.

Figure 5 presents a sample of position (x, y) and speed v evolution of a pedestrian recovered from his walking paths. These are the microscopic parameters of each individual pedestrian that were managed to be stored in .CSV files to characterize the group walking. These data are stored for each pedestrian detected along frames with respect to coordinate system of videos scaled to a size of 640 x 480 px, where lower left corner being coordinate $(0, 0)$ and frames are equivalent to the first quadrant of Cartesian coordinate system.

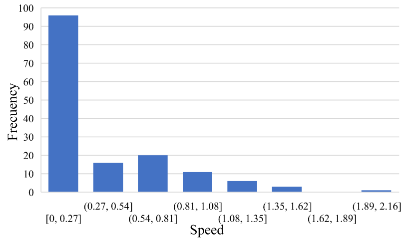
From positions and speed of pedestrians retrieved from real videos, histogram of speed distribution function shown in Figure 6 is obtained. As can be seen, crowd behavior at low densities in Figure 6.a shows a behavior similar



(a) Video at low density with $\mu = 1.34m/s$ and $\sigma = 0.67m/s$.



(b) Video at medium density with $\mu = 0.85m/s$ and $\sigma = 0.82m/s$.



(c) Video at high density with $\mu = 0.38m/s$ and $\sigma = 0.5m/s$.

Fig. 6. Frequency distribution of crowd average speed at low, medium, and high densities.

to the Gaussian function with mean $\mu = 1.34m/s$ and standard deviation $\sigma = 0.67m/s$ because pedestrians have few encounter conflicts with others. At medium densities, the behavior resembles a decreasing exponential function because present various encounters between pedestrians with $\mu = 0.85m/s$ and $\sigma = 0.82m/s$ (see Figure 6.b). At high densities, their behavior does not have a uniform behavior and most pedestrians have a concentrated speed at the lowest value between $0m/s$ and $0.27m/s$ with $\mu = 0.38m/s$ and $\sigma = 0.5m/s$ because crowd gatherings occur and pedestrians do not have enough space to maintain their social distance (see Figure 6.c).

5 Conclusions

This paper proposes a methodology for real-time video pedestrian tracking by applying an object detection algorithm based on convolutional neural networks. Walking paths between frames and some microscopic parameters of pedestrians are recovered to understand their evolution.

The main contribution of this work is description of walking paths and frequency distribution analysis of average speed to understand pedestrian crowd dynamics at different densities under regular walking conditions.

As future work, application of tracking algorithms capable of recognizing pedestrians occluded by obstacles and detecting multiple pedestrians in high densities is proposed to achieve more reliable results. In addition, the methodology can be extended to include other scenarios such as emergency exits or panic situations.

References

1. Aharon, N., Orfaig, R., Bobrovsky, B.-Z.: Bot-sort: Robust associations multi-pedestrian tracking. arXiv:2206.14651, (2022)
2. Bathija, A., Sharma, G.: Visual object detection and tracking using yolo and sort. International Journal of Engineering Research Technology, vol. 8, no. 11, pp. 705–708 (2019)
3. Camara, F., Bellotto, N., Cosar, S., Nathanael, D., Althoff, M., Wu, J., Ruenz, J., Dietrich, A., Fox, C. W.: Pedestrian models for autonomous driving part i: low-level models, from sensing to tracking. IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 10, pp. 6131–6151 (2020)
4. Dollár, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: A benchmark. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 304–311. IEEE (2009)
5. Feliciani, C., Murakami, H., Nishinari, K.: A universal function for capacity of bidirectional pedestrian streams: Filling the gaps in the literature. PloS one, vol. 13, no. 12, pp. 1–31 (2018)
6. Gawande, U., Hajari, K., Golhar, Y.: Pedestrian detection and tracking in video surveillance system: issues, comprehensive review, and challenges. Recent Trends in Computational Intelligence, pp. 1–24 (2020)
7. Murakami, H., Feliciani, C., Nishinari, K.: Lévy walk process in self-organization of pedestrian crowds. Journal of The Royal Society Interface, vol. 16, no. 153, pp. 1–10 (2019)
8. Pal, S. K., Pramanik, A., Maiti, J., Mitra, P.: Deep learning in multi-object detection and tracking: state of the art. Applied Intelligence, vol. 51, pp. 6400–6429 (2021)
9. Rangel-Huerta, A., Ballinas-Hernández, A., Muñoz-Meléndez, A.: An entropy model to measure heterogeneity of pedestrian crowds using self-propelled agents. Physica A: Statistical Mechanics and its Applications, vol. 473, pp. 213–224 (2017)
10. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 779–788 (2016)
11. Seer, S., Brändle, N., Ratti, C.: Kinects and human kinetics: A new approach for studying pedestrian behavior. Transportation research part C: emerging technologies, vol. 48, pp. 212–228 (2014)
12. Srazhdinova, A., Anvarov, S., et al.: Detection and tracking people in real-time with yolo object detector. Challenges of Science, vol. 2020, no. 3, pp. 69–75 (2020)
13. Stadler, D., Beyerer, J.: Improving multiple pedestrian tracking by track management and occlusion handling. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10958–10967 (2021)
14. Sundararaman, R., De Almeida Braga, C., Marchand, E., Pettre, J.: Tracking pedestrian heads in dense crowd. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3865–3875 (2021)
15. Xiao, Y., Zhou, K., Cui, G., Jia, L., Fang, Z., Yang, X., Xia, Q.: Deep learning for occluded and multi-scale pedestrian detection: A review. IET Image Processing, vol. 15, no. 2, pp. 286–301 (2021)

16. Zanlungo, F., Crociani, L., Yücel, Z., Kanda, T.: The effect of social groups on the dynamics of bi-directional pedestrian flow: A numerical study. In: Traffic and Granular Flow 2019. pp. 307–313. Springer (2020)
17. Zanlungo, F., Feliciani, C., Yücel, Z., Nishinari, K., Kanda, T.: Macroscopic and microscopic dynamics of a pedestrian cross-flow: Part II, modelling. Safety science, vol. 158, pp. 4–24 (2023)
18. Zuo, X., Li, J., Huang, J., Yang, F., Qiu, T., Jiang, Y.: Pedestrian detection based on one-stage yolo algorithm. Journal of Physics: Conference Series, vol. 1871, no. 1, pp. 1–7 (2021)