A COMPREHENSIVE REVIEW OF PEDESTRIAN TRAJECTORY EXTRACTION FROM VIDEO DATA: METHODS, APPLICATIONS, AND RESEARCH DIRECTIONS

Sabuj Biswas*[1], Mahdi Muntasir[2] and Md Asif Raihan[3]

[1] *Department of Civil Engineering, Bangladesh University of Engineering and Technology, Bangladesh, e-mail: sabuj.biswas.ce@gmail.com*

[2] *Department of Civil Engineering, Bangladesh University of Engineering and Technology, Bangladesh, e-mail: mahdiexplorer833@gmail.com*

[3] *Associate Professor, Accident Research Institute (ARI), BUET, Bangladesh, e-mail: raihan@ari.buet.ac.bd*

***Corresponding Author**

ABSTRACT

Pedestrian trajectory extraction from video plays an important role in many fields such as traffic safety, crowd management, autonomous driving, and urban planning. Researchers want to understand how pedestrians move, and they use video data because it captures natural behavior without disturbing people. Over the last two decades, improvements in cameras, computer vision, and machine learning have created many new methods for detecting and tracking pedestrians. Early studies use classical computer vision methods such as background subtraction and optical flow. Later studies use statistical models to represent pedestrian interaction. Recent studies use deep learning and multi-object tracking (MOT) to extract more accurate and stable trajectories, even in crowded scenes. A brief overview of existing research will be given in this study which includes research across indoor, outdoor, drone-based, mixed-traffic, and synthetic environments. This review is organized in both thematically and methodologically. Here classical computer-vision pipelines, model-based approaches, deep-learning detectors, multi-object tracking frameworks, and hybrid multi-sensor systems is covered.Also common challenges include occlusion, identity switching, dataset bias, and equipment limitations is also discussed. Finally, this study identifies emerging trends and also provides recommendations for future research that includes sensor fusion, domain adaptation, prediction-based tracking, and real-time scalable deployment.

Keywords: *Pedestrian, Trajectory, Video-Based Analysis Computer Vision, Trajectory Analysis*

1. INTRODUCTION

Pedestrians walk or move in complex ways in both indoor and outdoor environments. Study of these movements is important because they directly influence traffic safety, urban design, evacuation planning, and autonomous navigation systems. A pedestrian trajectory is a sequence of points that shows where a person walks over time. Researchers use these trajectories to measure walking speed, crowd density, human interactions, and collective movement patterns. These measures help researcher identify unsafe areas, design better facilities, and build accurate models of pedestrian behaviour. Traditional data collection methods, such as manual tracking or wearable sensors, often produce limited and incomplete information. These methods are slow, expensive, and unsuitable for large crowds or wide areas. Because of these limitations, video-based pedestrian trajectory extraction has become a popular and powerful tool. Video allows researchers to observe natural, uninterrupted behaviour, and cameras can cover large indoor and outdoor environments at low cost. Over the past twenty years, video sources have become more diverse and more advanced. Early studies mainly use fixed cameras in public spaces to track movement (Johansson et al., 2008). Later studies use stereo cameras, depth sensors, and multi-camera systems to solve problems such as occlusion and blind spots and to obtain more accurate tracking results (Hu et al., 2013; Liao et al., 2014). More recent research uses UAV, Infrared surveillance video, fisheye camera, or multi-view synchronized cameras or drone footage to observe large outdoor areas from above and reduce visual blockage from vehicles or obstacles (Bai et al., 2020; Shang et al., 2022; Yamane et al., 2025; L. Yang et al., 2020). During the same period, the methods used to extract trajectories have also evolved. Early studies mainly depend on classical computer vision techniques such as background subtraction (Fleuret et al., 2008), which work well in simple environments. But these techniques often fail in crowded, dynamic, or poorly lit scenes. Recent studies use deep-learning-based pedestrian detectors and identify people with higher accuracy in complex scenes (nighttime) (Zhao et al., 2024). Many studies also use multi-object tracking (MOT) frameworks to track each pedestrian across video frames and to maintain identity even when pedestrians overlap or occlude each other (H. Yang et al., 2024). This study reviews existing studies and covers a wide range of environments, including bottlenecks, indoor stations, outdoor intersections, aerial views, and synthetic environments. The study also describes the main methods used in those studies that include classical computer vision techniques, model-based approaches, deep learning detectors, multi-object tracking systems, and hybrid sensing methods. The goal of this paper is to present a clear and detailed academic summary. Also, highlights key techniques, major challenges, and potential research directions in pedestrian trajectory extraction. The paper organizes the studies in two ways. First, this paper groups the studies by scene type, including bottleneck experiments, indoor stations, outdoor mixed-traffic areas, UAV recordings, and synthetic environments. This thematic grouping helps explain how different environments create unique technical challenges. Second, the paper groups the studies by method type, including classical computer vision, statistical modelling, deep-learning-based detection, multi-object tracking, and hybrid multi-sensor systems. This methodological grouping shows how trajectory extraction technology has improved and diversified over time.

2. SCENE-BASED ANALYSIS

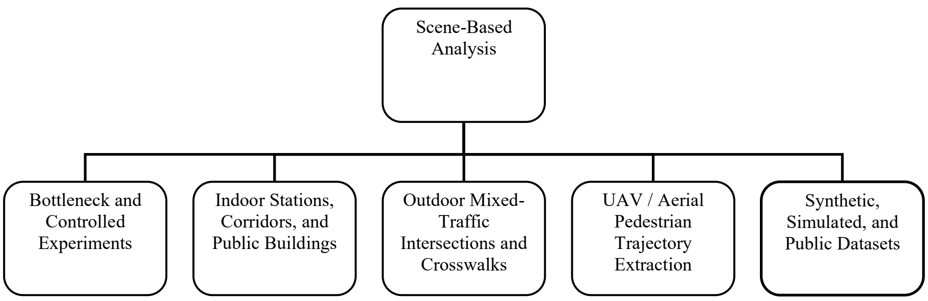

Figure 1: Scene Based Analysis

Bottleneck and Controlled Experiments

Bottleneck and controlled experiments help researchers to understand how pedestrians will behave in confined spaces where density and movement pressure increase. These settings help to understand lane formation, collision avoidance, and self-organizing patterns during crowd movement. Several past studies have examined this type of environment. Initially, Daamen and Hoogendoorn (2003) show how pedestrians move through a narrow entrance way in controlled conditions. The influence of density on flow through these narrow exits is investigated by Liao et al. (2014). Then, Zanlungo et al. (2023) analyse crossflow interactions and report the formation of temporary lanes. A high-precision trajectory is collected by Boltes and Seyfried (2010) from an overhead camera. Later, Plaue et al. (2012) and Plaue et al. (2011) used Controlled multi-camera setups, which are also used in indoor experiments. Moussaïd et al. (2009) examine how individual decisions produce collective patterns. Zhang and Seyfried (2015) show how boundary conditions shape corridor flow. Although these studies produce clean, high-quality trajectory data, their controlled conditions may not fully represent real-world complexity.

Indoor Stations, Corridors, and Public Buildings

Indoor spaces such as stations, university corridors, lobbies, and other public buildings usually have well-defined pedestrian routes. However, the movement is often difficult to capture because people overlap and block each other from view. A number of studies have tried to handle these issues. Hu et al. (2013) use a multi-camera system to track pedestrians in crowded indoor environments. Yang et al. (2020) propose a method to remove blind spots and improve tracking at high density. Lorgna et al. (2024) look at how people interact and avoid collisions inside public buildings. Other indoor work, including studies by Shang et al. (2022) and Yang et al. (2024), focuses on avoidance behaviour and pedestrians turn when different streams meet. Noh et al. (2022) combine indoor and outdoor cameras so that identities are not lost when people move between spaces. Although indoor environments benefit from consistent lighting and fixed camera positions, problems such as occlusion and high crowding still make trajectory extraction difficult.

Outdoor Mixed-Traffic Intersections and Crosswalks

Outdoor scenes are more complicated because pedestrians share space with vehicles, bicycles, and other moving elements, and conditions can change from one moment to the next. Lighting, weather, and clutter in the background all add to the challenge. Several studies deal with these situations. Belhadi et al. (2021) apply deep-learning methods to examine group behaviour and detect unusual movement outdoors, while Rodriguez et al. (2014) study pedestrian motion in natural outdoor settings. Zhang (2023) and Yang et al. (2024) study busy sidewalks and intersections data where pedestrians and vehicles move together. Bai et al. (2020) and Cosar et al. (2017) focus on how people and vehicles behave and interact in these areas. Yamane et al. (2025) look at safety signs and near-miss events at intersections to understand possible risks. Because outdoor environments involve heavy occlusion, shadows, and unpredictable movement, trajectory extraction generally depends on strong detection models and reliable multi-object tracking.

UAV / Aerial Pedestrian Trajectory Extraction

The UAV and drone videos give a useful top-down perspective that helps to reduce occlusion and capture large scenes. Aerial views help researchers to study and analyze crowd behavior in parks, plazas, and urban intersections. Shang et al. (2022) use UAV video from the VisDrone dataset and show that top-down aerial views help detect and track pedestrians even in busy urban scenes, but small pedestrian size and occlusion remain challenging. Dufour et al. (2025) used a high-altitude fixed "LargeView" camera overlooking a festival square, which provides wide-area coverage and enables the extraction of large-scale crowd dynamics. Instead of UAVs, Yamane et al. (2025) also focus on multi-view fixed-camera systems rather than drones, combining synchronized ground cameras to improve identity matching across viewpoints. Although UAVs have strong coverage, there is still challenges such as small pedestrian size, camera motion, and weather effects limit.

Synthetic, Simulated, and Public Datasets

Although synthetic datasets do not involve video extraction, they are included as a scene category because they provide trajectory data for benchmarking and model evaluation. Synthetic and simulated datasets allow researchers to test algorithms under controlled and repeatable conditions. Several studies in the dataset use simulation-based data.

Johansson et al. (2007) build physics-based models to understand how pedestrians move and interact. Saadat et al. (2012) use computer-made (synthetic) environments to test tracking methods and try different movement situations. Moussaïd et al. (2009) simulate crowd behaviour to see how simple choices by individuals can create large, group-level patterns. More recent studies, like Dufour et al. (2025), mix simulations with aerial videos to build hybrid datasets. Synthetic datasets are useful because they provide perfect ground-truth data and can be repeated many times. However, they still cannot match real-life complexity, such as changing lighting, mixed traffic, and unpredictable human actions.

3. METHOD-BASED ANALYSIS

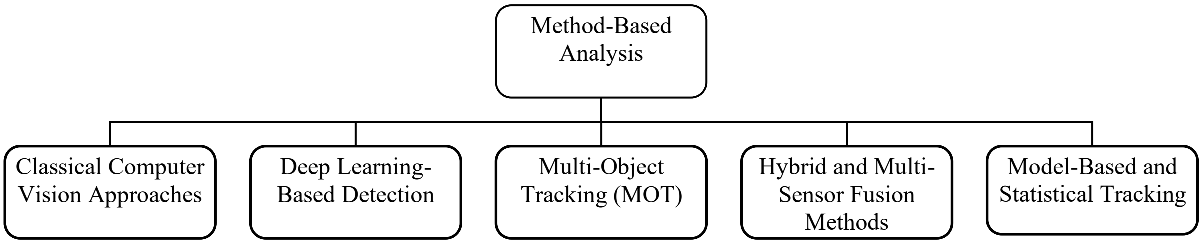

Figure 2: Method Based Analysis

Figure 2 presents the categories used in the method-based analysis. The following section provides detailed descriptions of each methodological group.

Classical Computer Vision Approaches

Classical computer vision methods were the first tools used for extracting pedestrian trajectories. These methods depend on simple pixel-level operations such as background subtraction, frame differencing, optical flow, blob detection, and Kalman filtering. They work well when the scene is simple, lighting is steady, and there are few people blocking each other. For example, Boltes and Seyfried (2010) use background subtraction and blob tracking to collect accurate trajectories in controlled lab settings. Daamen and Hoogendoorn (2003) also use classical image-processing methods to study how people move through bottlenecks. Zhang and Seyfried (2015) apply these traditional tools to examine how corridor boundaries affect pedestrian motion, and Rodriguez et al. (2014) use classical detection and tracking to study outdoor crowds. Saadat et al. (2012) test classical tracking methods inside simulation environments, while Cosar et al. (2017) combine simple motion features with classical detection to classify pedestrian behaviour outdoors. Although classical methods are fast and easy to use, they do not perform well in busy, noisy, or mixed-traffic environments. Shadows, occlusion, and changing backgrounds make tracking less accurate. For these reasons, most recent studies now use deep learning and multi-object tracking systems.

Deep Learning-Based Detection

Deep learning has changed pedestrian detection by using neural networks that learn patterns directly from data, instead of depending on hand-made features. Many recent studies use models like YOLO, Faster R-CNN, RetinaNet, and Mask R-CNN to improve accuracy in crowded or difficult scenes. Yang et al. (2020) use deep-learning-based detection to extract trajectories in crowded indoor spaces without leaving blind spots while Belhadi et al. (2021) use deep learning to study group movement and find unusual behaviour in outdoor urban areas . For detect pedestrians at night Zhao et al. (2024)

apply improved deep-learning model and shows that these models can work even in low-light conditions. Bai et al. (2020) and Yang et al. (2024) also use deep-learning detectors to study pedestrian movement in outdoor and indoor environments. Deep-learning detectors perform much better than classical methods in busy and dynamic scenes, but they also need large training datasets, strong computers, and careful tuning to work well in different places.

Multi-Object Tracking (MOT)

Multi-object tracking (MOT) is mainly used to follow pedestrians across video frames and create smooth, continuous movement paths. Most MOT systems combine a deep-learning detector with a matching method such as Kalman filtering, the Hungarian algorithm, or appearance-based re-identification. This helps reduce identity switches and keeps each person's track consistent, even in crowded or mixed-traffic scenes. Many past studies use MOT in different settings. Noh et al. (2022) use Mask R-CNN with Kalman filtering to track both pedestrians and vehicles at urban crosswalks. This allows them to study risky interactions using indicators like the Pedestrian Safety Margin. Yang et al. (2024) pair YOLO-based detection with MOT to track people in busy outdoor areas and still get reliable trajectories even when the crowd is dense. Bai et al. (2020) also use a YOLO–DeepSORT approach and show that it works well in low-light infrared videos, helping identify unusual behaviors such as standing still or hovering. UAV-based studies by Dufour et al. (2025) use PeTrack and MOT-style tracking to maintain stable pedestrian trajectories in crowded festival environments. Yang et al. (2020) apply MOT to fisheye indoor video to maintain consistent head tracking without blind spots. More recent multi-camera work from Yamane et al. (2025) improves MOT performance by using both motion and appearance features to match people across synchronized camera views. Although MOT models greatly improve identity stability and produce more complete trajectories, they still struggle in situations with strong occlusion, very dense crowds, major lighting changes, or when different people look very similar.

Hybrid and Multi-Sensor Fusion Methods

Hybrid and multi-sensor fusion methods combine data from multiple cameras or sensing systems to solve problems like occlusion, limited viewpoints, and low-light conditions. Several studies in the table use these approaches in both controlled and real-world environments. Initially, synchronized multi-camera setups were used by Plaue et al. (2012) and Plaue et al. (2011) in an indoor lobby to significantly remove blind spots and build complete pedestrian trajectories through calibration and view matching techniques. Later, Hu et al. (2013) also combined corner-mounted camera views using a probabilistic occupancy map to maintain identity consistency during occlusion in cluttered indoor areas. In a Recent Study, Dufour et al. (2025) combined fixed ground cameras with a high-altitude LargeView camera to improve tracking in dense nighttime festival crowds, and Yamane et al. (2025) used synchronized multi-view inputs to produce Bird's Eye View representations for stable trajectory extraction under heavy occlusion. Analyze pedestrian behavior in a complex urban square, Lorgna et al. (2024) merge video and geo-referenced scene data. These hybrid systems often provide better accuracy and more complete trajectories than single-camera methods. Although they require careful calibration, synchronization, and high computational resources. Despite this, they are increasingly adopted because they offer more reliable tracking in complex environments..

Model-Based and Statistical Tracking

Statistical and behavior-based approaches are not direct trajectory extraction methods because they do not take raw video frames and convert them into trajectories. Instead, these approaches work after trajectories have already been extracted or rely on simulated movement data. They help understand how individuals interact, how pedestrians avoid collisions, and how collective patterns such as lanes or stripes emerge in crowds. For example, Johansson et al. (2007) use interaction-based statistical models to analyze how pedestrians adjust their speed and direction based on neighbors. Moussaïd et al. (2009) show that simple local decision rules can produce large-scale collective patterns such as lane formation. Zanlungo et al. (2023) analyze microscopic and macroscopic dynamics in cross-flow experiments and show how temporary self-organized lanes appear when densities increase. These studies provide valuable insight into pedestrian behavior but do not detect or track pedestrians in video. Instead, they function as post-processing or analytical tools that help interpret trajectory datasets produced by computer-vision-based extraction systems.

4. CROSS-STUDY COMPARISON AND KEY FINDINGS

A comparison of existing studies shows several clear strengths and limitations in current research on pedestrian trajectory extraction from video data. One common finding is that overhead or top-view cameras perform very well in high-density pedestrian environments. This is mainly because these viewpoints reduce occlusion and make individual pedestrians easier to see. Studies by Boltes and Seyfried (2010), Daamen and Hoogendoorn (2003), and Zanlungo et al. (2023) show that ceiling-mounted or top-view cameras can produce accurate and stable pedestrian trajectories, especially in indoor or semi-controlled environments. Another important trend is the increasing use of deep-learning-based detection and tracking methods. Compared to traditional computer-vision techniques, deep learning models work better in difficult conditions such as nighttime scenes, crowded areas, and complex backgrounds. For example, Belhadi et al. (2021) use deep learning to analyze group-level pedestrian behavior, and Zhao et al. (2024) show that improved deep models can still perform reasonably well under low-light conditions. These results indicate that deep learning methods are more robust to changes in appearance and lighting. Studies that use aerial or high-altitude cameras also report good performance. Shang et al. (2022) show that UAV footage can cover large areas and reduce ground-level occlusion. Similarly, Dufour et al. (2025) use very high fixed cameras to record dense festival crowds and extract a large number of pedestrian trajectories at the same time. These studies highlight that camera placement is often as important as the tracking method itself. In addition to single-sensor approaches, several studies apply hybrid and multi-sensor fusion methods. These methods combine multiple cameras, different viewpoints, or different sensors (such as RGB and depth) to improve tracking reliability. Multi-camera approaches help maintain pedestrian identity during occlusion (Plaue et al., 2011, 2012), while Overhead stereo-vision and depth-based sensing approaches provide accurate spatial information in indoor environments (Boltes and Seyfried, 2013; Corbetta et al., 2014). Recent multi-view trajectory-level fusion methods, such as Yamane et al. (2025), further show that combining information across views improves robustness in complex scenes.

Despite these strengths, several key limitations still exist. Occlusion and identity loss remain major problems, especially in dense crowds and mixed-traffic scenes where pedestrians and vehicles interact. Some classical computer-vision methods still require manual correction after automatic tracking, as reported by Boltes and Seyfried (2010). Although deep-learning-based methods reduce the need for manual work, they usually require powerful hardware and may lose accuracy in very dark or visually complex scenes, as shown by Zhao et al. (2024). Another major challenge is poor generalization. Many methods perform well only in the type of scene they were trained or tested on, such as indoor corridors or pedestrian-only areas. When applied to new environments with different lighting, camera angles, or crowd behavior, their performance often decreases. Overall, the reviewed studies show strong progress in pedestrian detection and trajectory extraction. However, no single method works best in all situations. Practical use still requires better balance between accuracy, robustness, computational cost, and adaptability to different environments. Table 1 shows summary of scene types, methods, and limitations.

Table 1: Summary of Scene Types, Methods, and Limitations

Scene Type	Common Methods Used	Main Strengths	Main Limitations
Indoor, pedestrian only (corridors, stations)	Classical CV, depth sensors, PeTrack, multi-camera fusion	High accuracy, stable tracking	Limited area, occlusion in dense flow
Outdoor, pedestrian-	Deep learning + MOT,	Works well under	Identity loss and high

only (plazas, events)		changing lighting	computational cost in dense crowds
Mixed traffic (crosswalks, intersections)	YOLO-DeepSORT, FairMOT	Handles pedestrian vehicle interaction	Occlusion, identity switches, scene dependent calibration
Aerial / high-altitude views	Detection + tracking	Large area coverage, less occlusion	Small object detection, legal limits
Nighttime / low-light scenes	Enhanced deep learning models	Improved performance over classical CV	High computation, noise sensitivity

5. DISCUSSION

Over the past two decades, Pedestrian trajectory research using video data has developed. Early studies mainly used classical computer-vision tools such as background subtraction, optical flow, and simple tracking. Some works, such as Johansson et al. (2007), calibrated the Social Force Model using observed trajectories to understand pedestrian interaction patterns using real video data. Later, deep-learning approaches became more common. For example, Belhadi et al. (2021) used deep learning to analyse group behaviour, and Zhao et al. (2024) improved nighttime pedestrian detection and tracking by enhancing the YOLOP and DeepSORT pipeline. Some studies also use overhead or high-altitude video sources to reduce occlusion and provide wider coverage. Shang et al. (2022) showed that UAV footage reduces occlusion and captures larger areas, and Dufour et al. (2025) used a very high fixed camera to extract thousands of trajectories in a crowded festival setting. Different environments need different techniques. Indoor areas benefit from multi-camera fusion to remove blind spots (Plaue et al., 2011, 2012). Outdoor mixed-traffic scenes require strong detectors and stable tracking because of vehicles, shadows, and lighting changes (Noh et al., 2022). Aerial videos need accurate small-object detection since pedestrians appear very small in drone or tower-view footage (Shang et al., 2022). Despite progress, challenges remain. Many models do not generalize well to new locations, low-light scenes are still difficult, and mixed traffic often causes occlusion and identity loss. Deep-learning systems also need high computing power, which limits real-time use in low-resource settings. More efficient and robust methods are needed for practical deployment.

6. FUTURE RESEARCH DIRECTIONS

Future research on pedestrian trajectory extraction should focus on solving the main problems found in current studies. First, multi-camera and multi-sensor systems should be used more often. These systems can reduce blind spots, handle occlusion better, and improve tracking accuracy. Previous work by Plaue et al. (2011, 2012) demonstrate the benefits of multi-camera configurations. Second, models need better generalization ability so they can work well in different environments. Future methods should be able to handle changes in lighting, crowd density, and camera position. Techniques such as domain adaptation, training on multiple datasets, and environment-independent features can help improve performance in new scenes. Third, simulation and synthetic data should be used more widely. Simulated environments can represent rare or dangerous situations, such as evacuations or extremely dense crowds, which are difficult to capture in real life. Studies like Moussaïd et al. (2009) show that simulation is useful for testing and improving pedestrian models. Fourth, there is a need for lightweight enhanced deep-learning models that can run in real time on low-cost or edge devices. Reducing computational requirements while keeping good accuracy is important for real-world deployment. Finally, future systems should not only extract trajectories but also link them to safety and risk analysis. Connecting pedestrian trajectories with measures such as near-miss events or conflict indicators, as done by Yamane et al. (2025), can greatly increase the practical value of these systems for traffic safety and urban planning. Together, these research directions can make pedestrian-tracking systems more robust, efficient, and practical for real-world and smart-city applications.

7. CONCLUSION

Study on pedestrian trajectory extraction has improved a lot over the years. Early studies used simple image-processing methods, but newer work relies on deep learning and sometimes combines data from multiple sensors. These studies show that accurate trajectories are important for understanding how people move, improving safety, and supporting better urban planning. Different environments such as bottlenecks, indoor stations, outdoor intersections, drone videos, and simulated settings need different technical approaches. Deep-learning detectors and multi-object tracking work well in busy or complex scenes, while high or aerial views help reduce occlusion and capture larger areas. Even so, challenges like occlusion, identity loss, low-light conditions, and poor generalization across locations still cause problems. The discussion and future directions suggest clear ways forward: using multiple sensors to reduce blind spots, improving models so they adapt better to new places, developing lightweight systems that run in real time, and linking trajectory extraction with safety analysis. Overall, the progress is strong, but more research is needed to make pedestrian-tracking systems accurate, efficient, and reliable for real-world use.

8. DECLARATION OF USE OF AI

The authors acknowledge the use of artificial intelligence (AI) tools in the preparation of this manuscript. AI tools (ChatGPT 5.1) were used solely for minor language refinement, including improvements in grammar, clarity, and readability. All intellectual content, including the research design, data collection, data analysis, interpretation of results, and conclusions, was developed and validated entirely by the authors. All AI-assisted text was reviewed, edited, and verified to ensure accuracy, originality, and compliance with scholarly and ethical standards. No AI tools were used to generate scientific content or to conduct any part of the research.

9. REFERENCES

Bai, C., Gong, Y., & Cao, X. (2020). Pedestrian Tracking and Trajectory Analysis for Security Monitoring. 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), 1203–1208. https://doi.org/10.1109/ITOEC49072.2020.9141825

Belhadi, A., Djenouri, Y., Srivastava, G., Djenouri, D., Lin, J. C.-W., & Fortino, G. (2021). Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection. Information Fusion, 65, 13–20. https://doi.org/10.1016/j.inffus.2020.08.003

Boltes, M., & Seyfried, A. (2013). Collecting pedestrian trajectories. Neurocomputing, 100, 127–133. https://doi.org/10.1016/j.neucom.2012.01.036

Boltes, M., Seyfried, A., Steffen, B., & Schadschneider, A. (2010). Automatic Extraction of Pedestrian Trajectories from Video Recordings. In W. W. F. Klingsch, C. Rogsch, A. Schadschneider, & M. Schreckenberg (Eds.), Pedestrian and Evacuation Dynamics 2008 (pp. 43–54). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04504-23

Corbetta, A., Bruno, L., Muntean, A., & Toschi, F. (2014). High statistics measurements of pedestrian dynamics. arXiv. https://doi.org/10.1016/j.trpro.2014.09.013

Cosar, S., Donatiello, G., Bogorny, V., Garate, C., Alvares, L. O., & Bremond, F. (2017). Toward Abnormal Trajectory and Event Detection in Video Surveillance. IEEE Transactions on Circuits and Systems for Video Technology, 27(3), 683–695.https://doi.org/10.1109/TCSVT.2016.2589859

Daamen, W., & Hoogendoorn, S. P. (2003). Experimental Research of Pedestrian Walking Behavior. Transportation Research Record: Journal of the Transportation Research Board, 1828(1), 20–30. https://doi.org/10.3141/1828-03

Dufour, O., Dang, H.-T., Cordes, J., Korbmacher, R., Chraibi, M., Gaudou, B., Nicolas, A., & Tordeux, A. (2025). Dense Crowd Dynamics and Pedestrian Trajectories: A Multiscale Field Dataset from the Festival of Lights in Lyon. Scientific Data, 12(1), 718. https://doi.org/10.1038/s41597-025-04732-3

Fleuret, F., Berclaz, J., Lengagne, R., & Fua, P. (2008). Multicamera People Tracking with a Probabilistic Occupancy Map. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2), 267–282. https://doi.org/10.1109/TPAMI.2007.1174

Hu, T., Mutlu, S., & Lanz, O. (2013). Multicamera People Tracking Using a Locus-based Probabilistic Occupancy Map. In A. Petrosino (Ed.), Image Analysis and Processing – ICIAP 2013 (Vol. 8157, pp. 693–702). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41184-7_70

Jalili, M., Hakimpour, F., & Van Der Spek, S. C. (2018). Extraction of Usage Patterns for Land-Use Types by Pedestrian Trajectory Analysis. In M. R. Luaces & F. Karimipour (Eds.), Web and Wireless Geographical Information Systems (Vol. 10819, pp. 61–76). Springer International Publishing. https://doi.org/10.1007/978-3-319-90053-7_7

Johansson, A., Helbing, D., & Shukla, P. (2008). Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. https://doi.org/10.48550/ARXIV.0810.4587

Liao, W., Seyfried, A., Zhang, J., Boltes, M., Zheng, X., & Zhao, Y. (2014). Experimental Study on Pedestrian Flow through Wide Bottleneck. Transportation Research Procedia, 2, 26–33. https://doi.org/10.1016/j.trpro.2014.09.005

Lorgna, L., Ceccarelli, G., Gorrini, A., & Ciavotta, M. (2024). Video Analytics for Understanding Pedestrian Mobility Patterns in Public Spaces: The Case of Milan. Collective Dynamics, 9, 1–9. https://doi.org/10.17815/CD.2024.172

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

Noh, B., Park, H., Lee, S., & Nam, S.-H. (2022). Vision-Based Pedestrian's Crossing Risky Behavior Extraction and Analysis for Intelligent Mobility Safety System. Sensors, 22(9), 3451. https://doi.org/10.3390/s22093451

Plaue, M., Chen, M., Bärwolff, G., & Schwandt, H. (2011). Trajectory Extraction and Density Analysis of Intersecting Pedestrian Flows from Video Recordings. In U. Stilla, F. Rottensteiner, H. Mayer, B. Jutzi, & M. Butenuth (Eds.), Photogrammetric Image Analysis (Vol. 6952, pp. 285–296). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-24393-6_24

Plaue, M., Chen, M., Bärwolff, G., & Schwandt, H. (2012). Multi-View Extraction of Dynamic Pedestrian Density Fields. Photogrammetrie - Fernerkundung - Geoinformation, 2012(5), 547–555. https://doi.org/10.1127/1432-8364/2012/0138

Saadat, S., Teknomo, K., & Fernandez, P. (2012). Automation of Tracking Trajectories in a Crowded Situation. Fire Technology, 48(1), 73–90. https://doi.org/10.1007/s10694-010-0174-9

Shang, P., Zhou, X., & Hu, J. (2022). Pedestrian Trajectory Extraction Method Based on UAV Video. 2022 IEEE 8th International Conference on Computer and Communications (ICCC), 2081–2085. https://doi.org/10.1109/ICCC56324.2022.10065641

Yamane, T., Masumura, R., Suzuki, S., & Orihashi, S. (2025). MVTrajecter: Multi-View Pedestrian Tracking with Trajectory Motion Cost and Trajectory Appearance Cost (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2509.01157

Yang, H., Tang, Y., Yu, W., Li, X., & Zhang, P. (2024). Robust pedestrian tracking in video sequences using an improved STF module. Complex & Intelligent Systems, 10(1), 1365–1374. https://doi.org/10.1007/s40747-023-01224-y

Yang, L., Hu, G., Song, Y., Li, G., & Xie, L. (2020). Intelligent video analysis: A Pedestrian trajectory extraction method for the whole indoor space without blind areas. Computer Vision and Image Understanding, 196, 102968. https://doi.org/10.1016/j.cviu.2020.102968

Zanlungo, F., Feliciani, C., Yücel, Z., Nishinari, K., & Kanda, T. (2023). Macroscopic and microscopic dynamics of a pedestrian crossflow: Part I, experimental analysis. Safety Science, 158, 105953. https://doi.org/10.1016/j.ssci.2022.105953

Zhang, J., & Seyfried, A. (2014). Experimental studies of pedestrian flows under different boundary conditions. 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 542–547. https://doi.org/10.1109/ITSC.2014.6957746

Zhang, Q. (2023). Multi-object trajectory extraction based on YOLOv3-DeepSort for pedestrian-vehicle interaction behavior analysis at non-signalized intersections. Multimedia Tools and Applications, 82(10), 15223–15245. https://doi.org/10.1007/s11042-022-13805-z

Zhao, W., Ren, C., & Tan, A. (2024). Study on Nighttime Pedestrian Trajectory-Tracking from the Perspective of Driving Blind Spots. Electronics, 13(17), 3460. https://doi.org/10.3390/electronics13173460