A study on roaming behaviour of crowd in public space with the analysis in computer vision and Agent-based simulation

Yuansheng Wu[1], Zipei Zhang[2], Shimada Daii[3], Changlin Liu[4], Shaoyu Niu[5], Lipeng Xia[6], Liangfa Xiao[7] and Yasushi Ikeda[8]

[1, 2, 3, 4, 5, 6, 7] Keio University SFC, Tokyo, Japan
wuys@keio.jpl[1]
[8] The University of Tokyo, Tokyo, Japan

Abstract: People-flow in densely populated modern cities is a non-negligible factor to consider during urban space design. However, in the early phase of design, architects mainly deal with the static states of the human body. Other factors, such as duration and environment, should be considered when analysing crowd behaviour. The common methods for design presentation are drawings, sketches, models, etc. Dynamic visualisation of pedestrian behaviour might help architects have a better understanding of the design performance. Agent-based simulation has been explored by many researchers. Most of their pedestrian models have planned routes with origin and destination. Thus, we would like to propose a pedestrian model embedded with roaming behaviour that reflects the decision-changing process when exploring unfamiliar places. Previously, computer vision and agent-based simulation were different research streams. This paper discusses the initiative to study crowd behaviour using computer vision and agent-based simulation and examined the accuracy and validity of agent-based simulation by comparing the results from computer vision. Overall, this research aims to improve the accuracy of the agent-based simulation by utilizing the data extracted from surveillance video.

Keywords: Agent-based Simulation; Computer Vision; Crowd Behaviour; Data Visualisation.

1. Introduction

The purpose of architectural design is to create comfort for people and solve issues in a common shared space for communities. People-flow in densely populated modern cities is a non-negligible factor to consider during space design. However, in urban space design, architects mainly deal with the static states of the human body. The common methods for design presentation are drawings, sketches, models, etc. Spatial performance is yet to be well considered or presented during the design process. (Vroman and Lagrange, 2017) Architecture, which shapes the built environment, should understand and deal with people-flow. This research discusses the initiative to study crowd behaviour using computer vision and

Architectural Science and User Experience: How can Design Enhance the Quality of Life: 55th International Conference of the Architectural Science Association *2022,* P. Izadpanahi and F. Perugia (eds.), *pp. 96–105.*

agent-based simulation and to examine the accuracy and validity of agent-based simulation by comparing the results from computer vision.

An agent-based model is utilised to simulate people-flow through multiple design scenarios. Specifically, spatial performance can be evaluated during the early design phase, allowing designers to predict and visualise the results. Previously, many researchers and designers widely used and conducted multi-agent simulations. The simulation parameters were mostly decided by intuition, figures listed in literature reviews, and laboratory experiments. (Davidich and Köster, 2013) The impact on the crowds might be distinct in variable environments and scenarios, resulting in the uniqueness of the agent-based simulations. Furthermore, the spontaneous movements of people are influenced by their surroundings, such as signs and traffic signals. Dynamic pedestrian behaviour influenced by social interaction is yet to be deeply explored by previous studies.

In this research, we established two agent-based simulations with the software PedSim Pro and our proposed pedestrian behaviour model. For comparison and analysis, we extracted the data output from the selected surveillance video with Yolov5 and Deepsort algorithm. In the pedestrian simulations, the environment was set up based on the actual scene in the surveillance video. The location data of agents was recorded over a constant period, and heatmaps were drawn to reflect the density and distribution of agents during the time frame. However, it was difficult to consider the site influence in the simulations. By using computer vision, we analysed the surveillance video to extract the empirical data, which would make it comparable to the pedestrian simulations by data visualisation.

Overall, this paper focuses on an ongoing project exploring roaming behaviour model which aims to get more precise predictions based on the data learned from the analysis. Besides, we will apply the model to urban space design by testing the simulation using the generative design method at a later stage. Specifically, this paper demonstrates the efforts to help improve the time, cost, speed, efficiency, and accuracy of multi-agent simulations.

2. Literature Reviews

The behaviour of pedestrians in a specific environment is complex and changeable. The uncertain factors in the environment may affect the pedestrians' route choice, walking speed, etc. Although it is challenging to simulate behaviour due to the randomness of pedestrians, many researchers have developed different models to simulate pedestrian behaviour under various criteria and environments. To solve this problem, the Markov-Chain model is a common principle for explaining the uncertainty of behaviour analysis. Still, it does not consider the impact of previous results on the next prediction. (Henry *et al*., n.d.) In addition to the Markov-Chain model, Karoji *et al*. (2019) applied recurrent neural network (RNN) as a fundamental theory to develop their algorithm and simulated pedestrian behaviour in the Shinkiba station, Tokyo, Japan. They developed an algorithm based on the visual context of the pedestrian. By identifying the recognition degree of objects by agents, the influence of external factors on pedestrian behaviour in route choice from a starting point to a destination point could be evaluated. From the perspective of architectural design, an accurate pedestrian behaviour simulation can be an ideal tool for assessing the performance of the building under specific situations.

An agent-based simulation can be used for simulating individuals' or groups' decision-making, behaviour, and interaction in a particular environment. Previously, many scholars have applied pedestrian behaviour models to evacuation simulation. Helbing *et al*. (2000) created a computer model to simulate the evacuation performance of crowds based on social psychology and the relationship between panic behaviour and the structure of buildings. More specifically, they discussed the effect of placing building structures (columns) in front of the escape exits on evacuation performance. They created two scenarios: 1. There were no impediments in front of the exit 2. A column was placed in front of the exit. Unexpectedly, they pointed out that the probability of injuries is lower when the column is placed, which is counterintuitive. Later, Camillen *et al*. (2009) used Netlogo, an agent-based simulation developed by Java, and selected the Castro Ursino Museum (a castle in Catania, Italy) as the research object to simulate the visits to the museum rooms under normal conditions and analyse the pedestrian behaviour regarding the use of exits under emergency conditions. In addition to evacuation analysis, we expect to assess the spatial performance in terms of pedestrian behaviours after completing the project in the fields of architectural design and urban design.

The settings in an agent-based simulation are essential. The parameters and scenarios set by users might lead to inaccurate results. Some other researchers adopt the data from literature reviews and laboratory experiments. However, the research has a specific context, which might not suit the simulation perfectly. Besides, the empirical data collection process through experiments and on-site visits can be time-consuming. Thus, we are committed to finding a new method to evaluate the accuracy of the results by comparing the actual pedestrian data with the agent-based simulation. To obtain pedestrian data, Wu (2021) applied a deep learning method to visualise actual pedestrian trajectories. He first used an unmanned aerial vehicle (UAV) to record videos of three squares at Tianjin University, China, and then he used the YOLO, a deep learning object detection algorithm, to identify moving people in the video, from which pedestrian coordinate data was derived. The simulation of the actual pedestrian trajectory was presented by establishing heat maps and distribution maps. However, UAV usage is limited to wide-open outdoor spaces. Our research initiative is to avoid using UAVs but rather publicly available surveillance videos of public spaces because of their wide availability, low cost, and avoidance of privacy infringement.

Pan (2021) used PedSim Pro as an analysis tool to redesign the office plan according to the restrictions raised by the outbreak of COVID-19, such as social distancing. They simulated pedestrian behaviour in the office under three different scenarios. As a result, it could easily visualise the pedestrian trajectories and flows in different areas. However, in Pan's study, most of the building users are office workers who are familiar with the space and have a firm intention of reaching their destinations. Therefore, the behaviour model tends to be destination-oriented, which differs from the behaviour of visitors going into an unfamiliar place and exploring undecided destinations.

3. Methodology

The below figure shows the research methodology (Figure 1). Specifically, surveillance video was analysed using a deep learning-based algorithm. Parallelly, we set up the 3D model of the actual scene in the video by using Rhinoceros. The Grasshopper-based plug-in PedSim Pro and our proposed navigation model were utilised as pedestrian simulation tools. As for data output, we created heatmaps and boxplots for the data visualisation. A comparison was conducted to examine the difference between pedestrian simulations and data extracted from the surveillance video.

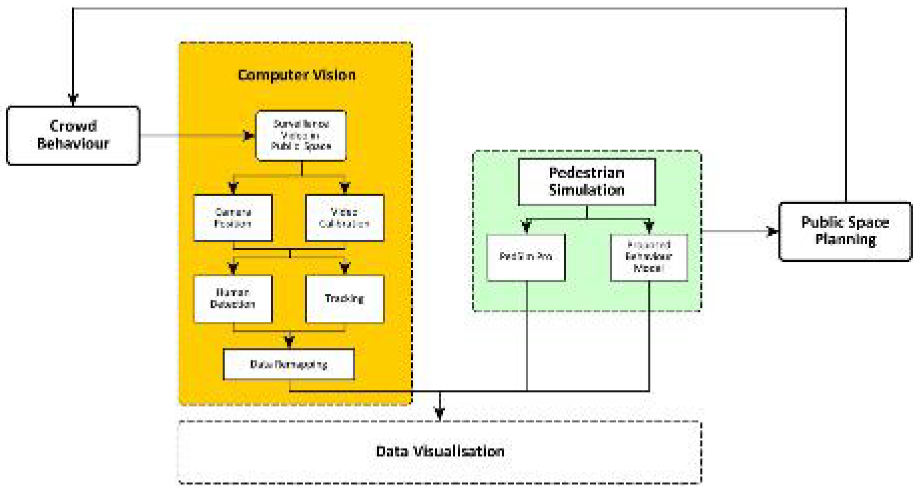

Figure 16: Research Methodology

3.1. Surveillance Video Selection

The surveillance video needs to satisfy the following criteria:

- Public space or large-scale architecture.
- Decent people-flow appears.
- Continuous live camera streaming.
- The resolution is high enough for object detection.
- It contains architectural elements.
- It does not constitute an infringement of privacy or copyright.

Figure 17: Surveillance Video captured at Fukui, Japan.

Taking these factors into consideration, Fukui Happy Terrace (Figure 2), a commercial facility located in Fukui City, Japan, was selected for our study. Specifically, the major obstacles identified in the scene are columns, food trucks, benches, and tables. Together with the shops, they become the significant factors that affect route choice in this plaza.

3.2. Video Calibration

The public camera parameters remain unknown to us. Therefore, calibrations are necessary for us to obtain accurate results from detection. We calibrated the video by considering two aspects: camera lens distortion and the position of the camera in the 3D model. The figure below (Figure 3) shows the distortion caused by the camera lens. It can be seen that the column in the left image is curved.

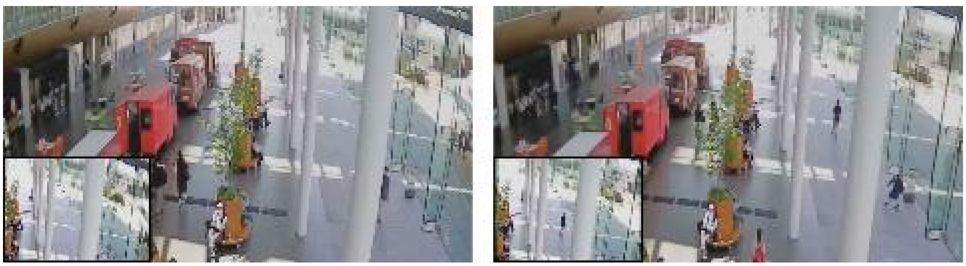

Figure 18: The distortion caused by surveillance camera.

A detailed floor plan was found on the official website of this facility, which has become the most reliable source for 3D modelling set up in Rhinoceros (Government, 2016). Besides, a simple version of the elevation drawing has been found on the building developer's official page (Facility, 2016). However, the exact positions of the food trucks, benches, and tables (Figure 2) are not included in the floor plan. First, according to the height provided by the elevation drawings, the rough location of the surveillance camera was found in the model. We adjusted the camera position in the model until the columns and roofs of the stores overlapped with the components in the actual scene. (Figure 4)

Figure 19: Model camera position calibration.

3.3. Crowd Detection

In our research, obtaining crowd behaviour data mainly relies on computer vision. Currently, we mainly focus on pedestrians' trajectories, while the actions of pedestrians will be combined in the future. For object detection, we applied YOLOv5 with Deepsort, an open-source algorithm, to detect and track pedestrians in the surveillance video. YOLOv5 is a library of object detection architectures with models pre-trained on the COCO dataset. DeepSORT is a deep-learning-based algorithm used for object tracking, assigning IDs to each object. By combining these algorithms, the movements could be tracked and recorded continuously.

Crowd movement data has been collected using YOLOv5 and DeepSORT algorithm. (Mikel, 2022) For a better detection result, we used CrowdHuman, a dataset focusing on crowd detection. Shao *et al*. (2018) created the weighted model trained based on 15,000 images with approximately 340,000 people, and the average number of people per image is around 23. Therefore, the dataset is suitable for the selected surveillance video.

Figure 20: Object detection and tracking by Yolov5 and DeepSORT.

The algorithms track each pedestrian's position and draw bounding boxes for each video frame. The frame IDs, person IDs, and coordinates are generated and exported to a text file. The data is then transferred to Excel for data sorting based on person IDs.

3.4. Remap the Location Data to the 3D Model

We remapped the location data from the images to the model through Rhino and Grasshopper to compare the results from human detection and simulations. Transforming the data into grasshopper, it was projected from the image to the model. The data has been filtered to reduce the errors resulting from human detection and tracking. We removed the coordinates data whenever the distance between two adjacent points is greater than 1 meter since the location is recorded every 0.5 seconds.

Figure 21: The sequence of camera calibration and data conversion.

3.5. Pedestrian Simulations

In the simulation of PedSim Pro, the agents move from the start gate to the destination gate along the shortest path. In the travelling process, the individual avoids other pedestrians and obstacles. Besides, the person profile is set to simulate pedestrians based on the actual travel speed. The origin and destination are fixed in the basic principle of PedSim Pro. If the interest points are added, the pedestrian trajectories will change due to the attraction and visibility of them. However, PedSim Pro does not consider route-changing behaviour. Pedestrians do not necessarily move towards the interest points or fixed destinations. Some pedestrians may wander aimlessly, and some pedestrians may always sit in their seats. This roaming behaviour is missing in the PedSim Pro. In our research, we aim to develop a pedestrian model that considers the route-changing behaviour when receiving different visual information from the surroundings.

Our proposed behaviour model aims to simulate the roaming behaviour of pedestrians. We developed an autonomous exploration pedestrian model driven by visual recordings of the points of interest. Agents can achieve events such as collision detection, information collection, avoiding others, and roaming exploration in an unfamiliar environment. In our model, agents are given starting points in advance, and multiple points of interest are set in the environmental configuration. When the simulation starts, the agent gains visual information from a random visible interesting point and walks to it. During the movement, the agent will collect information about the other points of interest based on vision, and the environmental information is saved in an activity list. To achieve this, multiple lines are connected between different interesting points and the agent. The line obstructed by obstacles will be removed. Based on the visual range of the agent set by users, the effective information can be sorted and saved. After a certain time, the agent will randomly select a location from the activity list, replace the previous interesting point, and move to it. After reaching the destination, the agent will stay for a while according to the settings and then continue to explore the environment.

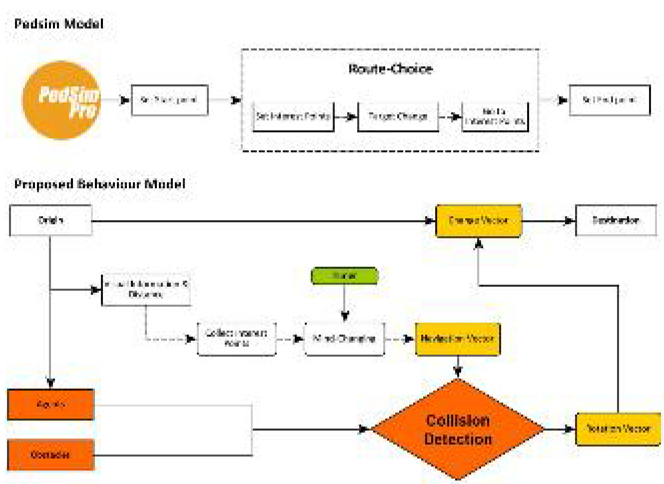

Figure 22: Diagrams of PedSim Pro (Top) and the proposed behaviour model (Bottom).

4. Result and Discussion

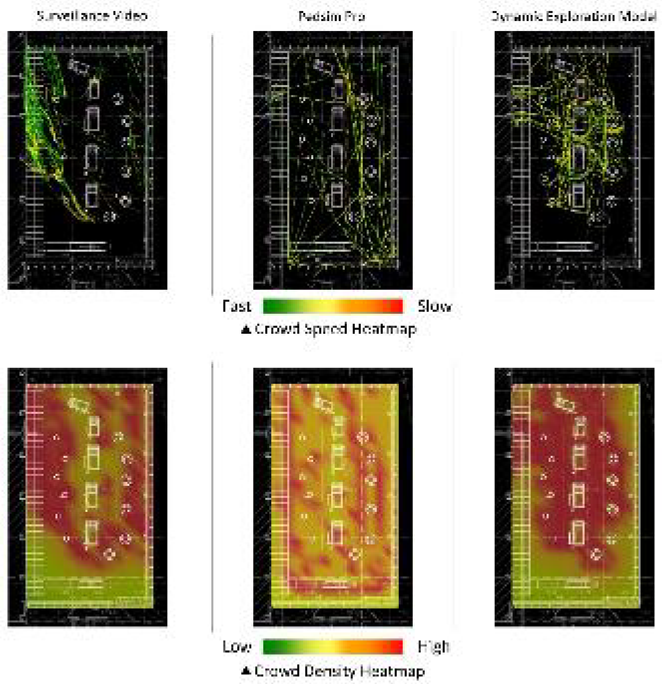

Figure 23: Result of data visualisation.

4.1. Discussion

Observing the surveillance video of our application, we can see that real-world pedestrians are passing through the site from unspecified places and the pedestrian trajectory shows a clear flow direction. In the PedSim Pro simulation, agents move from one or multiple points to the destination points. During the trip, agents have a certain probability of moving to interesting points. These origins, interesting points, and destination points are defined by users. The moving trajectories are composed of these points with obvious directionality. Our proposed behaviour model pre-sets one or multiple starting points and multiple points of interest. Though the agents are assigned destination points initially, as time passes, agents will change their original destinations with a certain probability set by the users during the travel due to the obtained information, resulting in an exploratory pedestrian simulation. The trajectories are woven into a grid through the starting point and a series of interesting points, which have obvious divergence.

From the comparison between the figures, in the real world, many factors may affect human behaviour. The PedSim Pro model focuses on the visible range of human sight in a plan view. However, our model discusses the visible range of human sight in three-dimensional space, the memory of information obtained while traveling, and the weights that have different preferences for points of interest. Crowd avoidance in reality is not as simple as volume collision detection in PC games. Through the surveillance video, it can be found that very few pedestrians have the behaviour of crossing the gap between the dining cars, while the PedSim model and our model show that a certain number of agents pass through the gap between the dining cars. In order to maintain adequate social distance and privacy, the actual avoidance behaviour may begin shortly after the obstacle or the opposite person comes into view. Some avoidance trajectories resemble a smooth tangent or sine function curve. Therefore, it might be possible to fit the results from the video by changing the parameters of obstacle avoidance in the next stage. By digitizing this behaviour pattern, it is possible to provide an accurate pedestrian behaviour model for architectural design and urban planning.

Crowd distribution, density, and types are different at different times of the day at the same location, such as peak hours and normal hours. The observed video clips concentrate on a specific period of the day. Therefore, the number and types of pedestrians passing through the site are relatively limited, which has certain limitations.

Overall, in the simulation performed by PedSim Pro, the necessity to set the origin and destination points leads to the agent's knowing the site information before departure. Therefore, this model is suitable for application to small-scale spaces such as shopping malls. In contrast to PedSim Pro, our proposed behaviour model is more inclined to explore unknown venues, such as applications to outdoor spaces such as exhibitions, amusement parks, and large commercial facilities. Recently, group phenomena have been increasingly studied in relation to the motions and dynamics of pedestrians. Sieben *et al*. (2017) state that crowds do not behave in an irrational and anti-social way. On the contrary, large groups of people usually move in quite an orderly and cooperative manner. The specific manifestation is that the individual's senses, cognition, judgment, and behavior appear in line with the behavior of most people. At the site we observed, the trajectory distribution of the crowd has obvious rules to follow, which may be due to the mutual reference of people on the site. There is an invisible force in the group that shapes the direction and distribution of the flow of people on the site. This consideration is still lacking in PedSim Pro and our proposed model.

4.2. Limitations

First of all, the video captured by only one surveillance camera meant that pedestrians could not be tracked completely throughout the building space, resulting in a missing portion of the pedestrian path in the lower left corner of the video. It was also difficult to collect richer data on pedestrian behaviour, as some popular destinations, such as access to other shopping streets and nearby bus stops, were out of this surveillance camera's field of view. Due to the limited camera range and the limitation that PedSim Pro can only simulate two-dimensional environments, it does not consider the movement of pedestrians on the second floor.

Second, the pedestrian trajectory and velocity map failed to show the complete trajectory of the pedestrian in the surveillance video. As mentioned earlier, YOLOv5 has some limitations. It cannot detect all human figures, especially when other objects block part or the whole of the human body. Random ID switching also occurs during the execution of the DeepSORT algorithm. Therefore, we had to remove some coordinate data that did not fit the pedestrian motion logic in the data cleaning, so discontinuous trajectory lines were generated. These trajectory lines cannot completely represent the actual pedestrian trajectory. In order to optimize the target detection, we may need to train with more images and find surveillance videos with higher image quality. We may also need to find another algorithm to pre-filter the data to optimize the trajectory map.

5. Summary and Future Research

This research proposes a new way to collect data by using surveillance video captured in public space and using the data to examine the accuracy of the pedestrian simulations. Compared to the previous research that collects data from laboratory experiments, field visits, or UAVs, the automated detection method can effectively collect data from the public space at a low time and labour cost. Furthermore, a pedestrian simulation considering roaming behaviour has been proposed. It is also suggested in this research that the destination-oriented pedestrian algorithm does not perfectly apply to many scenarios in public spaces.

We are in the middle stage of developing a comprehensive pedestrian behaviour model that considers wayfinding pedestrians affected by the visual information acquired from their surroundings. In the future, we will consider action recognition and define more parameters to improve the proposed behaviour model that can be applied to practical architectural design.

References

Camillen, F., Caprì, S., Garofalo, C., Ignaccolo, M., Inturri, G., Pluchino, A., Rapisarda, A. and Tudisco, S. (2009) *Multi agent simulation of pedestrian behavior in closed spatial environments*, ed.

Davidich, M. and Köster, G. (2013) *Predicting Pedestrian Flow: A Methodology and a Proof of Concept Based on Real-Life Data*, PLOS ONE, 8(12), e83355.

Facility, F. N. C. (2016) *Section drawings*. Available from: Fukui Nigiwai Commnucation Facility Happy Terrace <http://nigiwai.ftmo.co.jp/drawing/terrace/happiterrace_Sectionnal.pdf> (accessed 15 November 2021).

Government, F. C. (2016) *Plan drawings*. Available from: Fukui City Government <https://www.city.fukui.lg.jp/sisei/tkeikaku/sigaiti/nigiwai-bosyu_d/fil/zumen.pdf> (accessed 15 November 2021).

Helbing, D., Farkas, I. and Vicsek, T. (2000) *Simulating dynamical features of escape panic*, Nature, 407(6803), 487-490.

Henry, M., Worranat, P., Jeremy, J., Adrian, H., Christopher, W., Calvin, L. and Jimin, K. *Markov Chains*. (n.d.) Available from: BRILLIANT <https://brilliant.org/wiki/markov-chains/> (accessed 15 July 2022).

Karoji, G., Hotta, K., Hotta, A. and Ikeda, Y. (2019) Pedestrian Dynamic Behaviour Modeling - An application to commercial environment using RNN framework, *Proceedings of the 24th CAADRIA Conference*, Victoria University of Wellington, Wellington, New Zealand, 281-290.

Mikel, B. (2022) *Yolov5_StrongSORT_OSNet*. Available from: Gitbub <https://github.com/mikel-brostrom/Yolov5_StrongSORT_OSNet> (accessed 13 October 2021).

Pan, M. J. (2021) Redesigning the working space for social distancing: Modelling the movement in an open-plan office, *CIBSE Technical Symposium 2021*, 13-14 July 2021.

Shao, S., Zhao, Z., Li, B., Xiao, T., Yu, G., Zhang, X. and Sun, J. (2018) *Crowdhuman: A benchmark for detecting human in a crowd*, arXiv preprint arXiv:1805.00123.

Sieben, A., Schumann, J. and Seyfried, A. (2017) *Collective phenomena in crowds—Where pedestrian dynamics need social psychology*, PLOS ONE, 12(6), e0177328.

Vroman, L. and Lagrange, T. (2017) *Human movement in Public spaces: The use and development of motion-oriented design strategies*, The Design Journal, 20(sup1), S3252-S3261.

Wu, S. (2021) Approach to Auto-Recognition of Human Trajectory in Squares using Machine Learning-Based Methods - An application of the Yolo-v3 and the DeepSORT algorithm, *Proceedings of the 39th Education and research in Computer Aided Architectural Design in Europe eCAADe 2021*, University of Novi Sad, Novi Sad, Serbia, 361-370.